

THESIS / THÈSE

MASTER EN SCIENCES BIOLOGIQUES

Essai de reconstruction in vivo de mini-réseaux de régulation à partir des protéines E1A, E2F et RB

Drèze, Matija

Award date:
2003

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



**FACULTÉS UNIVERSITAIRES NOTRE-DAME DE LA PAIX
NAMUR**

Faculté des Sciences

**ESSAI DE RECONSTRUCTION *IN VIVO* DE MINI-RÉSEAUX DE
RÉGULATION À PARTIR DES PROTÉINES E1A, E2F ET RB**

**Mémoire présenté pour l'obtention du grade de
licencié en Sciences biologiques**

Matija DRÈZE

Juin 2003

Essai de reconstruction *in vivo* de mini-réseaux de régulation à partir des protéines E1A, E2F et RB

DRÈZE Matija

Résumé

La biologie du XX^e siècle a vu une accumulation de données biochimiques et génétiques relatives à la physiologie cellulaire en général. Ce mouvement s'est fortement accéléré au cours des dix dernières années grâce à l'essor de la génomique. La génomique fonctionnelle (physiogénomique) a pour but de donner sens à ces informations. L'interprétation des cartes d'interactions géniques et protéiques qui forment un réseau enchevêtré et complexe, dont on commence à disposer pour divers organismes modèles, est un défi majeur de la biologie. A priori, le problème devrait pouvoir être réduit à des modules fonctionnels, plus simples, au sein de ces réseaux intriqués.

réduction
v.s. non -

Nous proposons d'aborder la physiogénomique par une approche synthétique, modulaire et booléenne.

En pratique, nous avons entrepris de reconstruire *in vivo* deux réseaux de régulation protéines-protéines pour ensuite en éprouver le comportement.

Les réseaux d'intérêt répondent à des interactions dépendant de RB, E2F et E1A. L' "output" de ces réseaux reconstruits dans la levure prise comme une "éprouvette", peut être suivi par un "readout" tel l'activité d'un rapporteur GFP (Green Fluorescent Protein). Le premier réseau choisi voit E1A placé sous le contrôle d'un promoteur inductible par le cuivre (*Cup1*). Il s'agit d'un réseau "non-régulé" en ce sens que la présence du cuivre détermine l'état du système. Dans un second réseau, E1A est placé sous le contrôle d'un promoteur lui-même régulé par RB. Ce réseau présente une rétroaction négative unique et susceptible d'osciller.

Nous avons pu mettre en évidence une diminution de l'activité de RB lorsque E1A est mis sous le contrôle du promoteur *Cup1* en présence de cuivre (réseau non-régulé).

Veillée d'armes. Demain, mon mémoire doit être fini et entre les mains de mon promoteur. Je m'arrête un instant.

Mon mémoire a vraiment ressemblé à une aventure. Une aventure d'Astérix et Obélix. L'aventure est longue à raconter et ne fait que commencer.

Je présente les personnages. Obélix, c'est moi. Il est grand plutôt que gros. Astérix c'est Pierrot. C'est le plus petit des deux. Il est malin et rusé. Avec lui, Obélix est entre de bonnes mains. Abraracourcix, le chef et Panoramix, le druide, savent qu'avec lui, Obélix ne craint rien. Il est vraiment bien dans son rôle, une vraie nounou.

Abraracourcix justement, c'est le Professeur M. Vidal. C'est lui le chef qui n'attend pas que le ciel lui tombe sur la tête mais qui attend plutôt les résultats. Panoramix, le druide, c'est le sage qui réfléchit, qui a toujours un bon conseil à donner (ou un bon papier à proposer !). Il s'agit du Professeur J. Vandenhoute évidemment.

A vous trois, un tout grand merci. Pour l'accueil, les discussions, les conseils, et pour avoir choisi Pierrot comme tuteur. Il n'y avait pas mieux.

Mais le village ne se limite pas à ces quatre personnages. Il y a le compagnon d'Obélix, Idéfix. Oui Violaine, c'est toi. Cétautomatix, le forgeron qui est costaud. C'est Fred (ou gamin, fils de Lucien ! - c'est selon). Et bien sur Ordralphabétix, le poissonnier. C'est Delphine. A eux trois, merci pour les délires, les prises de bec, le soutien, les bons moments.

Merci au "Vidal-ome" (ou Vidal-um), qui n'est pas un camp romain retranché mais un village gaulois ami. Merci particulier à Mike, Muneesh, Monica, Tracey, Denis, Nicolas, Jeff, Bart, Chris et Siming. Merci pour les conseils, le matériel, les discussions.

Merci à Marcia, notre logeuse à Boston, grâce à qui notre séjour a été grandement facilité.

Merci à Frank du laboratoire du Dr P. Silver pour les plasmides.

Merci à Madame M. Raes, ainsi qu'à Messieurs D. Lambert, M. Salmon et J. Van Helden, membres du jury, de prêter attention à ce mémoire.

TABLE DES MATIÈRES.....	1
ABBREVIATIONS COURANTES	4
INDEX DES FIGURES	5
GLOSSAIRE	6

INTRODUCTION

1.BIOSYSTEMES MOLECULAIRES : FONDEMENTS CONCEPTUELS	9
1.1. LE PARADIGME DE JACOB ET MONOD : LA REGULATION BIOLOGIQUE PAR RESEAUX D'INTERACTIONS	9
1.2. LA PROPOSITION D'HARTWELL : LE MODULE, UNITE CONSTITUTIVE DES SYSTEMES ET RESEAUX BIOLOGIQUES.....	10
1.3. LA CONTRIBUTION DE THOMAS : L'ANALYSE LOGIQUE DE LA REGULATION EN RESEAU	11
1.3.1. Les systèmes à boucles	11
1.3.2. Un prototype de circuit "simple" : le switch du bactériophage lambda [30].....	16
2. INGENIERIE DIRECTE DE BIOSYSTEMES MOLECULAIRES.....	16
2.1. SYNTHÈSE DE SYSTEMES DE REGULATION <i>IN VIVO</i> : PRINCIPE ET FINALITE.....	16
2.2. EXEMPLE DE RECONSTITUTION FONCTIONNELLE D'UN MODULE TRANSCRIPTIONNEL : LE "REPRESSILATOR"	18
3. OBJECTIFS DE NOTRE TRAVAIL : LA RECONSTRUCTION D'UN MODULE DANS LA LEVURE	18
3.1. LES OUTILS, LES ACTEURS ET PARTENAIRES	18
3.1.1. Un outil nouvel outil de construction génétique : MultiSite Gateway™.....	19
3.1.2. Acteurs et partenaires.....	21
3.2. RESEAUX A RECONSTRUIRE.....	23
3.2.1. Réseau non-régulé.....	23
3.2.2. Réseau oscillant.....	24

RÉSULTATS

1. PRÉPARATION DES ÉLÉMENTS DE CONSTRUCTION DU VECTEUR D'EXPRESSION D'E1A	26
1.1. AMENAGEMENT DE LA MATRICE.....	26
1.1.1. Vecteur de départ et Stratégie d'ensemble du clonage	26
1.1.2. Etape 1 : Elimination du gène et du promoteur TRP1.....	28
1.1.3. Etape 2 : Clonage de la cassette MultiSite Gateway™ dans le vecteur intermédiaire pPC86MT1..	28
1.1.4. Etape 3 : Clonage de la cassette zéocine pour la création de pZACR42.....	30
1.2. CREATION DES MINI-BANQUES DE PROMOTEURS ET D'ORFS.....	32
1.2.1. Construction du mini-promoterome	33
1.2.2. Construction du mini-ORFéome.....	34
2. ASSEMBLAGE DES PIÈCES DU VECTEUR D'EXPRESSION D'E1A.....	35
3. RECONSTRUCTION ET TEST FONCTIONNEL <i>IN VIVO</i>	36
3.1. CHOIX DU "READOUT"	36
3.2. RESEAU NON-REGULE.....	36
3.2.1. Reconstruction dans la levure	36
3.2.2. Tests phénotypiques : révélation de la transactivation.....	36
3.2.3. Cinétique de mise en place des acteurs	39

DISCUSSION

1. CONSTRUCTION DU VECTEUR D'EXPRESSION D'E1A	40
2. RESEAU NON-REGULE.....	40
3. RESEAU OSCILLANT	41
3.1. RECONSTRUCTION DU RESEAU	41
3.2. CARACTERISATION PHENOTYPIQUE : RESULTATS DES TESTS DE CROISSANCE ET DES TESTS β -GAL....	41

CONCLUSIONS ET PERSPECTIVES

1. EXPERIENCES A FINALISER A PROPOS DES DEUX RESEAUX TESTES	43
1.1. LE MODULE-PILOTE NON-REGULE	43
1.2. LE RESEAU OSCILLANT	44
2. PERSPECTIVES DE RECONSTRUCTION A PLUS GRANDE ECHELLE	44
3. LA QUESTION DE LA PERTINENCE DE L'APPROCHE QUE NOUS AVONS SUIVIE.....	45

MATÉRIEL ET MÉTHODES

1. TECHNIQUES DE BASE DE BIOLOGIE MOLECULAIRE.....	46
1.1. UTILISATION DES ENZYMES DE RESTRICTION	46
1.2. FILL IN PAR LE FRAGMENT KLENOW DE LA DNA POLYMERASE I DES BRINS D'ADN PROTRUDANTS ..	46
1.3. DEPHOSPHORYLATION DES EXTREMITES 5'-PHOSPHATE DE BRINS D'ADN.....	46
1.4. LIGATION D'UN VECTEUR ET D'UN INSERT.....	46
1.5. REACTION EN CHAÎNE DE LA POLYMERASE (PCR)	47
1.5.1. Principe	47
1.5.2. Protocole de PCR (pour un volume final de 50 μ l)	47
1.5.3. Programme PCR	47
1.6. ELECTROPHORESE EN GEL D'AGAROSE	47
1.6.1. Principe	47
1.6.2. Réalisation d'un gel d'agarose 1%	48
1.7. PURIFICATION SUR GEL : QIAQUICK GEL EXTRACTION KIT PROTOCOL (AVEC UNE MICROCENTRIFUGEUSE).....	48
1.8. PURIFICATION SUR COLONNE : QIAQUICK PCR PURIFICATION KIT PROTOCOL (AVEC UNE MICROCENTRIFUGEUSE).....	49
1.9. MINIPREPARATION (MINIPREP) : QIAPREP SPIN MINIPREP KIT PROTOCOL (AVEC UTILISATION D'UNE MICROCENTRIFUGEUSE).....	49
2. TOPO CLONAGE (INVITROGEN).....	50
2.1. PRINCIPE	50
2.2. REALISATION PRATIQUE.....	50
3. REACTION DE MULTISITE GATEWAY™	51
3.1. REACTION DE BP CLONAGE.....	51
3.2. REACTION DE LR CLONAGE	51
4. TRANSFORMATIONS	52
4.1. TRANSFORMATIONS BACTERIENNES	52
4.2. TRANSFORMATION DE LEVURES	52

5. SOUCHES ET PLASMIDES	53
5.1. SOUCHES	53
5.1.1. Souches bactériennes (<i>E. coli</i>).....	53
5.1.2. Souche de levure.....	53
5.2. PLASMIDES (INVITROGEN)	53
5.2.1. <i>pPC86</i>	53
5.2.2. <i>pPICZalphaA</i>	53
5.2.3. <i>pCR®-BluntII-TOPO®</i>	53
5.2.4. <i>pDNR 201</i>	53
5.2.5. <i>pDNR P4-PIR</i>	53
6. MILIEUX	54
6.1. MILIEUX BACTERIENS	54
6.1.1. Milieu SOB (pour 1 litre).....	54
6.1.2. Milieu LB solide (Luria-Bertani).....	54
6.1.3. Milieu LB liquide (Luria-Bertani)	54
6.1.3. Milieu LB low-salt solide pour zéocine	54
6.2. MILIEUX LEVURES	55
6.2.1. Milieu non-sélectif.....	55
6.2.2. Milieu sélectif.....	55
7. TEST β-GAL.....	56

BIBLIOGRAPHIE

Abbréviations courantes

3AT	3-Triaminotriazole
5FOA	Acide 5-Fluorotique
AD	Domaine de transactivation
Amp	Ampicilline
ARS	Autonomous Replicating Sequence
CR2	Conserved region 2 (version courte)
DB	Domaine de liaison à l'ADN
DO	Densité optique
FL	Full length (version longue)
GFP	Green Fluorescent Protein
-His	Sans histidine
Kan	Kanamycine
LB	Luria-Bertani
-Leu	Sans leucine
M	Molaire
MCS	Site Multiple de Clonage
mg	milligramme
ml	millilitre
mM	millimolaire
NLS	Nuclear Localization Sequence
PCR	Réaction en chaîne de la polymérase
rpm	Rotation par minute
-Trp	Sans tryptophane
Zeo	Zéocine
β-Gal	β-Galactosidase
μl	microlitre

Index des figures

- Figure 1 : Représentation de l'opéron lactose et de la boucle positive de rétroaction
- Figure 2 : Exemple de cartes physiques protéine-protéine générée par la méthode du double hybride
- Figure 3 : Description des boucles de rétroaction
- Figure 4 : Illustration du destin d'une cellule au cours de sa différenciation
- Figure 5 : Mise en évidence d'une bistabilité comme comportement d'un réseau à rétroaction positive reconstruit en cellule
- Figure 6 : Représentation du switch du phage lambda
- Figure 7 : Le "*repressilator*"
- Figure 8 : Intégration et excision du bactériophage lambda dans le génome de *E. coli*
- Figure 9 : MultiSite Gateway™, réaction de BP Clonage
- Figure 10 : MultiSite Gateway™, réaction de LR Clonage
- Figure 11 : Illustration de l'importance de l'orientation des sites *att*
- Figure 12 : Principe du double-hybride en levure
- Figure 13 : Description de la souche MaV203
- Figure 14 : Action de E1A sur l'activité de RB
- Figure 15 : Interactions physiques entre DB-RB, AD-E2F et E1A
- Figure 16 : Représentation booléenne du réseau non-régulé
- Figure 17 : Représentation booléenne du réseau oscillant
- Figure 18 : Simulation *in silico* du comportement du réseau oscillant
- Figure 19 : Vue d'ensemble de la construction du vecteur de destination pZACR42
- Figure 20 : Carte de restriction du plasmide pPC86 digéré par *Sna*BI et *Bam*HI
- Figure 21 : Profils de restriction de pPC86 et de pPC86MT1
- Figure 22 : Vérification de la taille de pACR42 et de l'orientation de la cassette Gateway
- Figure 23 : PCR diagnostique de la présence de la cassette zéocine dans le vecteur TOPO
- Figure 24 : Test de fonctionnalité du vecteur de destination pZACR42 en levure
- Figure 25 : Restriction diagnostique de pZACR42
- Figure 26 : Vérification par PCR de la présence et de la taille de l'insert dans les vecteurs d'entrée après BP Clonage
- Figure 27 : Résultat de transformation de bactéries par les produits de LR Clonage
- Figure 28 : Vérification par PCR de la présence et de la taille de l'insert dans les vecteurs d'expression après LR Clonage
- Figure 29 : Test de transactivation des rapporteurs *URA3* et *HIS3* pour le réseau non-régulé
- Figure 30 : Test de transactivation du rapporteur *LacZ* pour le réseau non-régulé

- Boucle de rétrocontrôle
On parle de boucle de rétroaction lorsque les éléments d'un module ou d'un système sont fonctionnellement connectés de manière cyclique. Ces connexions sont réalisées de telle façon que le développement d'un processus est directement ou indirectement régulé par ses propres conséquences.
- Boucle de rétrocontrôle négative
Boucle de rétroaction dans laquelle chaque élément exerce une influence négative sur son propre développement ultérieur.
- Boucle de rétrocontrôle positive
Boucle de rétroaction dans laquelle chaque élément exerce une influence positive sur son propre développement ultérieur.
- Cassette Gateway
Séquence nucléotidique qui permet un recombinaison dans la technologie Gateway™. Elle est constituée de 5' en 3' par un site *att*, du gène de résistance au chloramphénicol Cm^R, du gène *ccdB* et d'un autre site *att*.
- Élément (ou variable)
Tout facteur en interaction physique ou fonctionnelle («logique») avec d'autres facteurs intervenant dans un module. A chaque élément est associé un état qui ne peut prendre que deux valeurs : 0 ou bien 1.
- Etat (d'un élément, d'un module, d'un système)
Statut (position) alternatif d'un élément, d'un module ou d'un système à un instant donné. Le nombre de valeurs qu'un état peut prendre est limité à deux : soit 0 soit 1 (la signification donnée aux valeurs peut être actif/inactif, présent/absent, ON/OFF...).
- Génomique
Domaine de la biologie qui analyse un organisme non plus au niveau d'un gène à la fois mais au niveau de tous ses gènes à la fois soit au niveau du génome. Ce domaine a généré des données en masse parce qu'il s'agit d'une analyse à grande échelle.
- Glossaire
Aide-mémoire.
- Ingénierie directe et reverse [1] [2]
Domaine de la physiogénomique qui cherche à reconstruire des réseaux d'interactions. L'ingénierie directe (bottom-up) est une approche de reconstruction, d'assemblage des pièces d'un module. On cherche à implémenter *in vivo* un module qui réponde à un comportement particulier. L'ingénierie reverse cherche à prédire le comportement d'un réseau à partir d'un modèle.
- Input
L'input (ou les inputs) est la donnée d'entrée d'un opérateur logique. Il est à la logique booléenne ce que la vue d'un prédateur est à une proie.
- Interactome
Carte des interactions protéines/protéines d'un organisme entier.

- Logique booléenne

La logique booléenne relève de la mathématique. Elle a été élaborée par le logicien et mathématicien anglais Georges Boole (1815-1864) et consiste à utiliser des signes d'opérations mathématiques pour établir un lien logique entre deux éléments distincts, c'est-à-dire de définir le type de relation existant entre ces deux éléments.

C'est un système de modélisation dans lequel les éléments ne prennent que deux valeurs, 0 ou bien 1, qui s'excluent mutuellement.

- Mémoire

Grands devoirs.

- Module

Ensemble de constituants (éléments) interagissant et intervenant dans un processus défini au bénéfice (sélectif) de l'organisme. Les modules, grâce à leurs interactions internes, gouvernent des fonctions discrètes ; celles-ci à leur tour seraient intégrables et "modulables" grâce à des interactions externes ou intermodulaires. Le module, constitue une entité fonctionnelle, entendant par là que le module *per se* a une fonction discrète mais susceptible d'être soumise à régulation.

- Opérateur logique

Lien logique (dans le contexte de la logique booléenne) entre plusieurs éléments. On distingue trois opérateurs (donc trois liens logiques): l'opérateur NON (encore appelé inverseur), l'opérateur ET (l'un et l'autre), l'opérateur OU (ou inclusif - soit l'un, soit l'autre, soit les deux).

- ORF

Séquence codante avérée ou hypothétique.

- ORFéome

Collection de séquences codantes confirmées (CDS) et prédites (pCDS). Dans notre cas, ces ORF's sont insérées dans des vecteurs compatibles avec la technologie du MultiSite Gateway™ (pENTRY-ORFéome).

- Output

L'output est la donnée de sortie d'un opérateur logique. Il est à la logique booléenne ce que la fuite est à la proie qui voit un prédateur, c'est à dire une réponse à un stimulus.

- Physiogénomique

Domaine de la biologie de l'ère post-génomique qui cherche à exploiter l'énorme masse de données générées par la génomique afin d'en trouver un sens physiologique.

- Promoterome

Collection de promoteurs. Dans notre cas, les promoteurs sont portés par des vecteurs compatibles avec la technologie du MultiSite Gateway™ (pENTRY-promoterome).

- Promoteur
Séquence nucléotidique, en amont d'une ORF, de taille souvent difficile à dé-limiter, dont une partie est reconnue par des facteurs de transcription qui peuvent activer la transcription de cette même ORF.
- Readout
Gène rapporteur placé en parallèle d'un module qui permet d'en suivre l'évolution au cours du temps.
- Règle de parité
La question du signe d'une boucle dépend uniquement de la parité du nombre d'interactions négatives dans la boucle: on parle de boucle négative si ce nombre est impair, on parle de boucle positive si ce nombre est pair.
- Réseau
Ensemble d'éléments, de modules, de systèmes en interactions.
- Système
Association de modules qui entretiennent des relations physiques ou fonctionnelles les uns avec les autres.
- Table des états
Table donnant pour toutes les combinaisons des états des éléments, l'état correspondant de la fonction.

1. Biosystèmes moléculaires : fondements conceptuels

1.1. Le paradigme de Jacob et Monod : la régulation biologique par réseaux d'interactions

Il y a plus de quarante ans, François Jacob et Jacques Monod proposaient le modèle de l'opéron lactose [3] [4]. Celui-ci décrit la régulation génétique: interaction entre une protéine de régulation (répresseur) et une séquence-cible (opérateur) qui met en position " OFF " l'expression de l'unité de transcription en aval. Le switch " ON " fait suite à l'interaction d'un inducteur avec la protéine-répresseur. Dans le cas particulier de l'opéron lactose décrit, l'inducteur (le lactose) est également substrat d'une enzyme codée par l'opéron. De la sorte, un élégant système de rétroaction positive, dont la logique répond à un avantage évolutif, est mis en place [Figure 1]. Bien que le modèle de l'opéron tel quel ne soit pas généralisable, l'idée que la cellule régule son fonctionnement non seulement par la régulation enzymatique abondamment étudiée par les biochimistes mais également par la régulation génétique selon un système de régulation dual opérateur-répresseur (cis-trans) et que de tels systèmes peuvent donner lieu à des " switches " a été une des plus fécondes de la biologie contemporaine.

Jacob et Monod avaient anticipé eux-mêmes que les mécanismes régulateurs, biochimiques et génétiques trouveraient de nombreuses illustrations sur le plan des mécanismes cellulaires du métabolisme, du développement et de la différenciation. Dans une publication [5] dont le titre seul explicite la clarté et la force prédictive de leur pensée, ils qualifient ces mécanismes de " téléonomique " et d' " intrinsèquement libres ". Ils entendent par là que la construction d'un système régulateur recèle une finalité et que celle-ci est exclusivement fonction de la manière dont le circuit de régulation est construit. Ils conçoivent, " construisent " théoriquement, plusieurs modèles (six) en se servant du type d'éléments de contrôle connus à ce moment dans la cellule (*E.coli*). Le jeu des connexions (" circuitry ") et des interconnexions imaginées *a priori* révèle une richesse de comportements, steady-state, oscillatoire, différenciateurs, de mémorisation...qui donnent à ces modèles une grande valeur heuristique pour la recherche biologique. Il est intéressant de remarquer que selon Jacob et Monod, la signalisation à la base de la régulation biologique reposait sur des réseaux de régulation dont la structure déterminait la fonction. La relation structure-fonction à laquelle la biologie nous a accoutumé revêt cependant ici un caractère imprévu. Les propriétés " émergentes " de la structure du réseau sont souvent hors de portée de l'intuition et nécessitent pour être appréhendées la mise en oeuvre de formalismes et de modèles simplificateurs [6-8]. Nous reviendrons sur le champ de recherches qui s'est ainsi ouvert à la biologie théorique dans le sillage des travaux sur la régulation des systèmes biologiques. À l'époque, les auteurs mesuraient bien que la mise à l'épreuve directe de leurs modèles était provisoirement hors de portée. Ils pensaient, à juste titre, qu'il faudrait attendre que les processus physiologiques et le développement soient mieux documentés au plan biochimique. Notons que l'hypothèse d'une ingénierie *a priori*, non sur papier, mais *in vivo* n'effleure pas les esprits les plus éclairés à l'époque. Nous reviendrons sur cette idée plus loin. Depuis les travaux de Jacob et Monod, il y a eu en effet une accumulation de données tant sur les régulations ou dérégulations génétiques que biochimiques [9]. Il semble clair que si on veut obtenir une interprétation fonctionnelle cohérente et intégrée de ces mécanismes dans la

réponse adaptative ou dans la différenciation, la modélisation s'avère une étape indispensable. C'est un objectif assigné à ce qu'on appelle la " biologie des systèmes¹ " [10, 11].

1.2. La proposition d'Hartwell : le module, unité constitutive des systèmes et réseaux biologiques

Depuis que Jacob et Monod ont publié leurs réflexions et modèles téléonomiques théoriques, on a vu une accumulation de données génétiques et biochimiques relatives à la réponse adaptative, la différenciation ou, de manière générale, à la physiologie cellulaire. Parmi les apports les plus marquants de ces recherches, il y a sans nul doute les voies de signalisations intra- et inter- cellulaires qui sont venues s'ajouter aux signalisations génétiques du type Jacob et Monod. Considérant l'image globale obtenue aujourd'hui du fonctionnement du vivant, on pourrait dire que la cellule possède des machineries, les unes faites de composants macromoléculaires assemblées en complexes individualisables physiquement (ribosome, transcriptosome, protéasome, spliceosome, enhanceosome, réplisome, ...) les autres n'ayant d'existence que par les relations fonctionnelles de leurs composants. Dans l'un comme dans l'autre cas, ces ensembles tirent leur fonctionnement intégré et leur identité fonctionnelle des relations ou interactions existant entre leurs composants. Mais ces fonctions ne peuvent être aisément prédites en étudiant les propriétés des composants isolément.

Dans un article séminal [12], Hartwell évoque la nécessaire transformation de la biologie moléculaire en biologie modulaire. Le concept de module recouvre les ensembles fonctionnels cités plus haut : un module comprend un ensemble de constituants (des éléments) interagissant et intervenant dans un processus défini au bénéfice (sélectif) de l'organisme. De tels modules, grâce à leurs interactions internes, gouvernent des fonctions discrètes ; celles-ci à leur tour seraient intégrables et " modulables " grâce à des interactions externes ou intermodulaires. On voit donc que le module, constitue une entité fonctionnelle [13], entendant par là que le module *per se* a une fonction discrète mais susceptible d'être soumise à régulation. Si dans certains cas la conception de l'entité modulaire est assez aisée (par exemple, le module ribosome constituant la tête de lecture suffisante pour fonctionner dans l'appareil de traduction, dans des conditions données), ce n'est pas toujours le cas (par exemple, les modules indistincts des voies de transduction du signal [14] présentant des recouvrement mutuels multiples). En principe, dans la mesure où le module est une entité remplissant une fonction (avantageuse) dans l'édifice cellulaire on peut imaginer qu'il est reconstituable et testable *in silico* dans un modèle théorique mais également dans une reconstitution physique. Ainsi, la mise en présence adéquate des composants d'un module supposé serait susceptible de reproduire, de voir émerger, la fonction.

L'importance et la nécessité de la conversion proposée par Hartwell de la biologie moléculaire, " *un gène à la fois* ", à la biologie modulaire sont particulièrement perceptibles dans le contexte postgénomique.

Si on considère la [Figure 2] de la carte présentant le réseau des interactions physiques pairées entre toutes les protéines d'un organisme modèle (*C. elegans*) [15], il est clair que la définition de modules discrets à l'intérieur de ce réseau dense est indispensable pour tenter de l'interpréter². De plus, l'établissement des influences entre eux s'impose si on veut extraire le sens biologique de ces données.

¹ Les systèmes sont des associations fonctionnelles d'éléments discrets. Le concept lui-même remonte à des travaux anciens.

² Certaines méthodes d'analyse de réseaux par clustering sont en cours de développement (Han et al. – « *Dynamic analysis of yeast interactome supports the notion of proteome modularity* » soumis pour publication – communication personnelle).

En résumé, la conception modulaire de la Biologie apparaît au centre de la physiogénomique et plus généralement de toute problématique de l'interprétation globale du fonctionnement des systèmes biologiques [16] [17]. Nous verrons comment la "*synthetic biology*" est une approche de cette interprétation.

1.3. La contribution de Thomas : l'analyse logique de la régulation en réseau

Nous avons tenté d'introduire, plus haut, le lecteur à la problématique des systèmes de régulation biologique telle que nous-même l'avons perçue au contact de certains travaux pionniers. Le concept de propriétés émergentes, inhérentes à la structure du réseau, était déjà énoncé par Jacob et Monod et la complexité potentielle des comportements des réseaux, était illustrée à partir de quelques exemples conçus *a priori*. Toutefois, la période qui a suivi fut caractérisée par l'énorme essor de toute l'analyse génétique et surtout la création de tout le corpus de la biologie moléculaire moderne. Le gène a été et reste l'unité analytique et physique de la biologie fonctionnelle. En effet, le succès de la biochimie, de la microbiologie et de la génétique moléculaire conjuguées était fondé sur le précepte "*one gene, one enzyme*" et la recherche par "*one gene at once*".

Ce n'est qu'en marge de ce vaste domaine de la Biologie moléculaire que l'étude pour elle-même des réseaux de régulation a vu le jour. Objet de travaux pionniers et véritablement remarquables, ceux-ci sont cependant demeurés relativement confinés aux domaines dit de la biologie "théorique". Ce terme, certainement révérencieux, cache mal le fait que la biologie "en marche" n'y plaçait pas son intérêt prioritaire. L'avènement très récent des approches postgénomiques qu'on pourrait qualifier, par contraste avec la biologie moléculaire classique, de "holistique" rend particulièrement évident, comme nous l'avons évoqué, le recours aux concepts de réseaux et modules. C'est dans le contexte de l'analyse des modules et réseaux biologiques complexes que les travaux pionniers retrouvent toute leur actualité. Le contenu des pages qui suivent est directement emprunté, parfois littéralement à R. Thomas et plus particulièrement à un texte paru à l'occasion du jubilé de l'Institut de Biologie moléculaire de Rhode Saint-Genèse (ULB). [18-22] [23]

1.3.1. Les systèmes à boucles

Malgré l'extraordinaire diversité des êtres vivants, il est une caractéristique qui paraît aussi fondamentale que le fait même de leur parenté : l'existence de boucles de rétroaction dans les systèmes de régulations génétiques. On parle de boucle de rétroaction lorsque les éléments d'un module ou d'un système sont fonctionnellement connectés de manière cyclique. Ces connexions sont réalisées de telle façon que le développement d'un processus est directement ou indirectement régulé par ses propres conséquences [Figure 3a].

Il en est ainsi si A influence B qui à son tour influence C qui lui-même influence A (A, B et C peuvent être des molécules qui agissent les unes sur les autres) [Figure 3b].

Dans une boucle de rétroaction simple, c'est-à-dire non-ramifiée, chaque élément est activé ou inhibé de manière directe uniquement par l'élément qui le précède immédiatement dans la boucle [Figure 3c et 3d].

Les connexions entre composants cellulaires impliqués dans un processus fonctionnel peuvent être neutres ou avoir un sens fonctionnel positif ou négatif. Ainsi, par exemple, on peut convenir qu'une kinase a un effet positif sur une cible alors que la phosphatase correspondante

aurait dans ce cas un effet négatif. Lorsqu'un réseau comporte plus d'une boucle de rétroaction, l'intuition de tous les jours n'est plus d'une grande aide pour prévoir le comportement de l'ensemble c'est-à-dire de la dynamique du réseau. Hors de portée donc de l'intuition, ces situations complexes requièrent pour être appréhendées la mise en oeuvre de formalismes et de modèles simplificateurs susceptibles de se prêter à des traitements informatiques.

1.3.1.1. Types de boucles et comportements

Dans une boucle de rétrocontrôle, chaque élément exerce une influence sur l'évolution de tous les éléments de la boucle, dont lui-même. Il y a des boucles dans lesquelles chaque élément exerce une influence positive sur son propre développement ultérieur : "boucle positive". En revanche, des boucles dans lesquelles chaque élément exerce une influence négative sur son propre développement ultérieur sont qualifiées de "boucle négative". La question du signe de la boucle dépend uniquement de la parité du nombre d'interactions négatives : on parle de boucle négative si ce nombre est impair, on parle de boucle positive si ce nombre est pair.

1°) Comportement des boucles négatives

Dans une boucle négative, chaque élément exerce une influence négative sur son propre développement via les autres éléments (s'il y en a) de la boucle - avec pour conséquence une homéostasie. Le comportement d'un thermostat illustre bien cette situation : la logique du circuit est telle qu'il tend à maintenir les éléments impliqués dans la boucle à une valeur située entre la limite inférieure à laquelle serait le système s'il était complètement OFF et la limite supérieure s'il était complètement ON, sans boucle de rétroaction.

2°) Comportement des boucles positives

Dire que chaque élément d'une boucle exerce un effet positif sur sa propre évolution, revient à dire que chaque élément tend à garder la valeur qu'il avait déjà auparavant : élevée si elle était déjà élevée, basse si elle était basse. Il s'agit d'un comportement autoentretenu. Si la présence d'un élément est la seule condition requise pour sa production, ou bien l'élément est absent, et il le reste indéfiniment, ou bien il est présent et il le reste indéfiniment (sauf contraintes externes) : cercle vicieux. Dans le cas d'une boucle à deux éléments antagonistes (chacun exerce une influence négative sur l'autre) le système peut se stabiliser dans deux états correspondant à la présence soit de l'un soit de l'autre [Figure 3e].

3°) Propriétés émergentes des systèmes à boucles : Homéostasie et différenciation

→ Homéostasie

Au plan biologique, le rôle essentiel des boucles de rétroaction négative est de générer l'homéostasie proche d'un état "steady state". Cette définition recouvre deux situations physiologiques différentes. Quand le "steady state" est stable, ou faiblement oscillatoire, le système fonctionne vraiment à la manière d'un thermostat et les oscillations se situant à l'intérieur d'une fourchette étroite peuvent être considérées comme des imperfections du système de stabilisation. Cependant, quand ceci conduit à des oscillations d'ampleur appréciable, celles-ci peuvent avoir une signification physiologique par elles-mêmes. Dans l'un comme dans l'autre cas, R. Thomas qualifie le comportement des boucles négatives de

" normatif " du fait que ces boucles maintiennent le système à (ou proche de) un seul état, une valeur, une norme.

→ Différenciation

R. Thomas mentionne que la première suggestion claire que des différences épigénétiques (non-liées au génome), c'est-à-dire une différenciation, répondaient en fait à la multistationnarité potentielle du système a été notée par Max Delbrück en 1949.

La vision épigénétique du développement, ou pourquoi les cellules se différencient en lignées distinctes de manière irréversible (" lineage commitment "), alors qu'elles possèdent le même génome, avait déjà été exprimée précédemment en termes imagés par Waddington [Figure 4]. Il comparait le destin d'une cellule à celui d'une balle dévalant les pentes de vallées aux profils intriqués ; une fois une vallée atteinte, le retour à la totipotence n'est plus possible. Ci-dessous, nous examinons le rôle des boucles positives dans l'état multistationnaire de systèmes avec pour conséquence que ceux-ci ont de ce fait plusieurs destins potentiels. [24, 25] [26]

Les boucles positives possèdent l'étonnante propriété de permettre un choix entre des situations stables distinctes : on dit que les boucles positives ont un caractère diversificateur, différenciatif alors que les boucles négatives, on l'a vu, ont un caractère "normatif". Un système à deux éléments, formant une boucle positive [Figure 3e] peut se représenter comme un paysage comportant deux bassins séparés par un col correspondant à un état instable. Ainsi, un tel système à trois états dont deux sont dits stables, et pour cette raison qualifiés du nom d'attracteurs c'est-à-dire de terminus possibles des trajectoires d'un système. [Figure 5] [27, 28].

La différenciation a un objectif profondément différent de l'homéostasie: au lieu de contraindre le système à rester constant autant que possible, elle permet un **choix**, une décision entre deux ou plusieurs possibilités. Comment deux cellules au patrimoine génétique identique peuvent-elles générer des lignées différentes c'est-à-dire être et transmettre stablement des différences épigénétiques? Les mécanismes responsables de ces décisions lors de la différenciation cellulaires sont l'objet d'innombrables recherches. R. Thomas part d'exemples "simples" pour illustrer la situation et proposer un raisonnement explicatif.

Considérant, par exemple, le cas déjà évoqué de l'opéron lactose [29], on peut établir des conditions expérimentales telles que des populations bactériennes à génome identique soient bloquées de manière durable dans l'un ou l'autre état (opéron lactose ON ou bien OFF) en fonction d'un " détail " de leur histoire. En effet, le lactose du milieu, s'il rentre dans *E. coli* induit l'opéron, dont le gène de la perméase qui " facilite " la rentrée du lactose. Dès lors, il est possible de séparer en deux une même culture en présence d'une concentration de départ en lactose capable d'induire le système. La première sous-culture (A) cependant est rapidement diluée de telle sorte que l'opéron n'est jamais induit (Ra : répresseur actif/perméase absente/lactose interne absent). A la seconde (B) on laisse le temps à la concentration de lactose d'induire (déréprimer) l'opéron (Ri : répresseur inactif/perméase présente/lactose rentrant) et donc la perméase ; ensuite on dilue le milieu jusqu'à une concentration de lactose qui, seule (sans perméase), ne permettrait plus l'induction initiale. On voit alors que A demeure toujours non induit (Ra) à la concentration de lactose " critique " mais, étonnamment que B, alors qu'elle se trouve à la même concentration de lactose " critique ", reste induit pour 150 générations ou plus ! L'état différenciatif acquis se

maintient épigénétiquement. L'explication réside dans le fait que l'opéron requiert la présence d'un inducteur (= dérépresseur) intracellulaire pour être actif (Ri : répresseur inactivé = opéron déréprimé) et donc assurer la synthèse d'une perméase. Mais, sous les conditions expérimentales décrites à concentration extracellulaire devenue critique, l'internalisation de l'inducteur requiert la perméase. Ainsi, seules les cellules qui ont expérimenté une concentration externe de lactose (avant dilution du milieu de culture) tel que celui-ci a diffusé dans la cellule pour atteindre un niveau interne déclenchant l'opéron et induit la synthèse d'un peu de perméase (état A : lactose interne/Ri/perméase active) maintiendront la rentrée de lactose grâce à l'activité perméase : on se trouve en présence d'un cercle vicieux..

Lors de l'interaction entre le bactériophage lambda et *E. coli* un " switch " s'opère entre lyse et lysogénie. Si cI, la protéine répresseur de lyse est active (situation A), le phage reste à l'état latent (lysogénique) dans le génome d'*E.coli*. C'est seulement si et seulement si un peu de protéine cro (Control of cI repressor and other things) apparaît (situation B) et réprime la production de cI que le système bascule vers la lyse. Une situation à logique similaire prévaut donc dans les deux cas. La situation A voit au départ la présence de A (Ra ou cI) qui exerce un effet négatif sur B (lactose interne ou cro), capable à son tour d'exercer un effet négatif sur A. C'est donc si A disparaît ou B apparaît que le système peut « choisir » de prendre et de rester dans l'autre état stable B, lequel sera transmis épigénétiquement. Le schéma de la [Figure 1] illustre la similitude logique qui prévaut : une boucle positive est à l'origine de la réussite du processus différenciatif. R. Thomas propose que " la présence d'une boucle positive dans la structure d'un système est une condition nécessaire à l'existence d'attracteurs multiples " et fait la conjecture que " tout processus de différenciation de nature épigénétique repose sur un mécanisme comportant une boucle positive ".

1.3.1.2. Traitement booléen

1°) Définition

Même les organismes les plus simples - les virus - contiennent dans leur génome l'information pour des mécanismes complexes de régulation. A fortiori cette complexité se retrouve-t-elle, encore accrue, chez les organismes supérieurs. L'analyse de tels systèmes de régulation devrait permettre de les décomposer en un nombre plus faible de types d'éléments de contrôle connectés correctement. En somme, comme le dit Thomas, il s'agit de découvrir le " squelette " du modèle et d'étudier ensuite les comportements que ce squelette, représentant la structure logique du système, permet de prévoir. Avec ceci en tête, il ne semble pas irréaliste d'aborder la physiologie des systèmes, même très complexes, par l'étude des circuits ou câblages de contrôle qu'ils comportent. Le squelette ou la structure d'un système sera représenté par un graphe qui est orienté, c'est-à-dire où les interactions pairées considérées entre éléments ou partenaires sont affectées de signes (+/-) qualifiant la nature de cette interaction (activation, inhibition) et par des indications qui précisent la nature logique de la connexion entre deux ou plusieurs opérations.

L'étude des phénomènes engendrés par les boucles a suscité beaucoup d'intérêt et a été sujette à un certain nombre d'études théoriques surtout basées sur des analyses continues. Cependant, les systèmes de régulations peuvent aussi être analysés en utilisant une approche dite logique (ou booléenne), dans laquelle chaque fonction et élément ne peut prendre que deux valeurs, typiquement 1 ou 0 (qui peuvent signifier oui/non, vrai/faux, présent/absent, actif/inactif, etc...) qui s'excluent mutuellement.

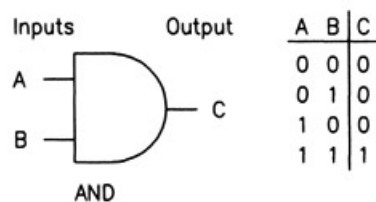
Nous allons ci-dessous montrer comment R. Thomas propose de traiter en termes booléens, des concepts, des situations et des modèles dans le domaine de la régulation génétique. Les objections à une approche binaire sont bien connues et évidentes. On peut dire que ce langage empêche l'expression de nuances, que les substances ne devraient pas être traitées comme présentes ou absentes, que la nature des processus est souvent continue, que l'ensemble des paramètres physico-chimiques doit être connu et inclus pour autoriser une modélisation pertinente et efficace des systèmes biologiques etc... Il y a certes un côté caricatural à ce type de description mais il s'agit de la contrepartie à sa capacité simplificatrice et heuristique. En réalité, la description théorique d'un système prend sa valeur s'il peut être mis à l'épreuve expérimentale. Nous reviendrons sur cette possibilité après en avoir explicité les préliminaires théoriques.

2°) Opérateurs logiques

Par opérateurs logiques, on entend une relation logique entre des éléments d'un module. On parle d'input(s) unique ou multiples (données d'entrée de l'opérateur logique) de "gate" (opérateurs logiques) et d'output unique (donnée de sortie de l'opérateur logique). L'état d'un module est révélé par l'état d'un gène rapporteur, le readout, qui est placé en parallèle du module d'intérêt. On verra l'application de ces concepts plus loin dans le texte. On distingue trois opérateurs logiques à partir desquels toutes les opérations plus complexes peuvent être réalisées (opérateurs NAND, NOR, XOR et XNOR). Par convention, les états des différents éléments sont exprimés en termes de 1 ou non 1.

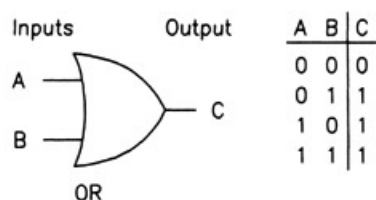
→ Opérateur ET (AND)

En logique booléenne, on parle de produit logique. La relation entre les inputs et l'output peut s'exprimer comme ceci: "**Si A ET B sont 1, ALORS C est 1**". En d'autres mots, la fonction logique C prend la valeur 1 si A et B ont tous les deux la valeur 1. A chaque opérateur, on associe une représentation symbolique ainsi qu'un tableau reprenant les états des différents éléments: la table des états (ou table de vérité).



→ Opérateur OU inclusif (OR)

En logique booléenne, on parle de somme logique. La relation entre les inputs et l'output peut s'exprimer comme ceci : "**Si A OU B est 1, ALORS C est 1**" (OU étant pris dans un sens inclusif). En d'autres termes, C prend la valeur 1 si l'une au moins des variables à la valeur 1. Voici la table des états ainsi que la représentation symbolique associée à l'opérateur OU.



→ Opérateur NON (NOT)

En logique booléenne, on parle du complément (inverseur). La relation entre l'input et l'output peut s'exprimer comme ceci : "**Si A est 1, ALORS C est NON 1**". Voici la table des états ainsi que la représentation symbolique associée à l'opérateur NON.

1.3.2. Un prototype de circuit "simple" : le switch du bactériophage lambda [30]

Un certain nombre de gènes bactériens et viraux sont impliqués dans la décision entre la lyse et la lysogénie chez les bactériophages tempérés dont il fût brièvement question plus haut. Plusieurs tentatives ont été réalisées pour modéliser ce réseau de régulation complexe. Il est possible de simplifier ce réseau jusqu'à ne plus tenir compte que des deux acteurs centraux, à savoir les protéines cI et cro. Pour appréhender le comportement du réseau dans son ensemble, il suffit de comprendre que :

- cI empêche la production de cro et favorise sa propre production
- cro empêche la production de cI

La représentation classique de ce réseau est à la [Figure 6a]. Ce réseau simplifié est essentiellement une boucle positive susceptible de conduire à un choix différenciatif entre deux situations. C'est pour cette raison que le réseau est appelé "switch". A la [Figure 6b], on résume les influences des deux acteurs l'un envers l'autre en termes booléens.

Il est d'emblée possible de percevoir la raison du choix offert au bactériophage : la présence de la boucle de rétroaction positive – sur laquelle se greffe une autre boucle de même type. Mentionnons que ce type d'approche a été appliqué, notamment par l'école de Bruxelles, à l'étude logique de divers autres switchs biologiques.

2. Ingénierie directe de biosystèmes moléculaires

2.1. Synthèse de systèmes de régulation *in vivo* : principe et finalité

On se rappellera les constructions de systèmes de régulation de Jacob et Monod dans le texte fondateur déjà cité. Les constructions étaient purement théoriques et considérées comme des exemples hypothétiques dont la réalité biologique pourrait être vérifiée éventuellement lorsque les chercheurs mettraient de tels réseaux à jour.

R. Thomas après avoir imaginé des systèmes purement abstraits, puis avoir analysé le comportement pouvant être engendré par le " squelette-logique " révélé au sein de quelques systèmes naturels, aborde explicitement la voie inverse. Il pose la question suivante : Comment construire un circuit qui satisfasse à des exigences préconçues ? C'est donc ce qu'il appelle une démarche " synthétique " ou autrement dit, inductive qui - au contraire de la précédente, analytique ou déductive - qui part de la logique pour en déduire les comportements, va cette fois de l'observation de comportements pour en induire le modèle logique approprié. Thomas note qu'en choisissant cette démarche inverse de celle de la

Biologie explicative usuelle il adopte la même démarche que celle de créateurs ou concepteurs de circuits logiques pour les équipements de contrôle industriel par exemple. On dirait aujourd'hui qu'il s'agit de faire de l'ingénierie directe ou de la synthèse de systèmes. Formellement, il est clair que plusieurs modèles de synthèse, différents, pourront satisfaire les mêmes réponses ou solutions comportementales. L'auteur développe, toujours au plan théorique, une série d'analyses dans ce cadre. Nous allons examiner ici la possibilité de passer à l'épreuve, non plus théorique mais pratique de circuits observés ou préconçus.

L'idée de vérifier la fonctionnalité de systèmes ou d'éléments de systèmes (modules) de régulation par synthèse *in vivo* était jusqu'il y a peu totalement irréaliste. Le clonage offre, au moins en principe, la possibilité de reconstituer, de construire au sens strict un réseau dans une cellule et d'en éprouver le comportement émergent. Cette approche "bottom-up" est par essence un moyen *ad hoc* pour comprendre un système de régulation et est susceptible de compléter utilement ce qu'on appelle la "*computational systems biology*" [31, 32]. A côté de l'ambition déclarée par certains chercheurs de chercher à reconstituer une cellule à partir d'un jeu nécessaire de gènes (Craig Venter) ou de tenter de créer des vésicules *ab initio* dans lesquelles on implémenterait les machineries d'expression et de réplication et auxquelles on fournirait l'apport d'énergie, il n'y a eu à ce jour que quelques tentatives concrètes de reconstitution et d'analyse de comportement de modules simples par transformation dans des cellules hôtes (*E. coli* ou *S. cerevisiae*).

De l'ensemble des données introductives que nous avons réunies jusqu'ici, il ressort que l'approche analytique et synthétique des processus régulatoires représente un défi d'importance pour tenter de "comprendre" les systèmes complexes que la biologie met à jour (voies de transduction, réseaux génétiques de génomes même simples comme le HIV) depuis des décennies et dont le paysage paraît de plus en plus inextricable. Dans ce contexte, on ne peut qu'être frappé par la masse des données dites de postgénomique ou de génomique fonctionnelle (physiogénomique) qui s'accumulent dans les banques de données. La publication de cartes d'interactions protéiques constitue un des domaines les plus actifs de la physiogénomique. Force est de constater que ces cartes sont encore descriptives et n'auront de "sens" physiologique – et ne constitueront donc à proprement parler des données physiogénomiques – que lorsqu'un "sens" sera donné aux interactions présentes dans le réseau. Il convient d'abord en effet d'affecter un signe (+/-) aux liens d'interactions paires établis dans ces cartes descriptives. Comme noté précédemment, une interaction peut être neutre ou dotée d'un signe. Nous avons vu l'importance de la règle de parité pour la qualification des boucles de rétroaction, positive ou négative éventuellement présentes dans le réseau. Les états différenciatifs – ou, au contraire, homéostatiques (ou normatifs) émergent des systèmes ou réseaux à boucle positive ou négative respectivement. L'observation simplement visuelle des cartes suggère que dans certaines régions à connectivité dense, il pourrait effectivement exister des circuits à boucles.

En réalité, la biologie d'aujourd'hui place le chercheur devant un nouveau défi. La collection de données va à un rythme inégalé et la clarté du paysage s'en ressent. Avec la séquence génomique en main on a eu coutume de dire, voire de croire – qu'on avait le "*blue print*" de l'organisme. On s'aperçoit que le nombre de caractères (altérables/phénotypiques) de l'organisme est beaucoup plus élevé que celui des gènes supposés y pourvoir. Le fait qu'il y ait une combinatoire des produits de gènes en interaction est une clé de la complexité phénotypique.

Si on veut aborder la complexité phénotypique avec quelque chance de succès nous pensons qu'on doit se référer au paradigme de Jacob et Monod sur l'existence des réseaux de

régulation, aux outils d'étude logique des réseaux selon Thomas et enfin à la proposition de Hartwell sur l'existence dans la cellule d'entités fonctionnelles modulaires. Les atlas d'interactions en cours d'établissement pour divers organismes modèles (et les "omes" et "omics" correspondants) sont à considérer probablement comme les vrais "blue prints" ou autre "pierre de Rosette" qu'il s'agit de déchiffrer pour obtenir "l'intelligence" du fonctionnement intégré de la cellule.

Une des moyens d'accéder à l'intelligence des cartes serait donc d'y reconnaître et d'y analyser des modules "élémentaires" selon l'approche usuelle de "bombing" de liens ou de nœuds dans les réseaux... Complémentaire à l'approche analytique, la synthèse ou la reconstitution et l'observation comportementale *in vivo* de modules définis prend toute sa signification dans le contexte de la post-génomique ou physiogénomique. Nous examinerons ci-après un exemple d'étude entreprise à une échelle modeste dans cette voie pionnière.

2.2. Exemple de reconstitution fonctionnelle d'un module transcriptionnel : le "*repressilator*"

L'exemple choisi veut illustrer la faisabilité d'une approche synthétique des mécanismes de régulation par réseaux. Il a fait l'objet, avec quelques autres, d'une revue de synthèse parue récemment dans un insert spécial du journal Nature (Nature insight, nov. 2002). Les auteurs de ce remarquable travail ont désigné et construit, dans *E. coli*, un réseau synthétique en vue d'obtenir un comportement particulier. Le système comporte trois répresseurs transcriptionnels en un réseau appelé *repressilator*. Le système induit de manière périodique la synthèse de GFP comme "readout" de son état dans les cellules individuelles. Les oscillations résultantes, typiquement d'une périodicité de plusieurs heures, sont plus lentes que le cycle de division cellulaire. Pour cette raison, l'état de l'oscillateur doit être transmis de manière épigénétique, de génération à génération [Figure 7a et 7b].

Cet exemple montre la faisabilité d'une approche synthétique des mécanismes de régulation et illustre la valeur de la méthodologie pour disséquer des comportements attendus (ou inattendus) de tels systèmes. Notons que, jusqu'à présent, le nombre d'éléments des modules reconstitués reste extrêmement réduit et qu'il s'agit, à une exception près de régulations transcriptionnelles.

3. Objectifs de notre travail : la reconstruction d'un module dans la levure

Notre projet s'inscrit dans le contexte de la "biologie des systèmes". Il vise spécifiquement à reconstruire *in vivo* un module de régulation simple, c'est-à-dire à un nombre faible d'acteurs dont on sait les relations fonctionnelles. Le module choisi répond à des interactions protéines-protéines dépendant de RB, E2F et E1A. Ces protéines sont décrites comme intervenant dans la régulation du cycle cellulaire (en sens divers). Les protéines seront produites par expression dans la levure comme outil. La technique du MultiSite Gateway™ est exploitée pour les clonages sériés. Le "monitoring" d'un rapporteur ("readout") permettra de suivre "l'output" du module dans des conditions variées et de vérifier sa capacité à produire un comportement désiré.

3.1. Les outils, les acteurs et partenaires

3.1.1. Un outil nouvel outil de construction génétique : MultiSite Gateway™

En 1961, Jacob et Monod pressentaient la nécessité de pouvoir reconstruire un réseau dont la théorie prédisait le comportement. Mais ils constatent avec regret l'absence d'informations expérimentales. Le système Multisite Gateway™ nous permet de reconstruire un réseau simple d'interactions de protéines. Il s'agit bien là d'une approche d'ingénierie moléculaire: à partir d'éléments, nous cherchons à reconstruire un module.

3.1.1.1. Principe biologique sous-jacent

Le principe est dérivé de l'observation de l'intégration-excision du génome du bactériophage lambda. Comme les autres virus, il peut tuer la cellule qu'il infecte (voie lytique). Mais il peut également former une association stable (voie lysogénique) " symbiotique " avec son hôte : l'ADN du phage s'intègre au génome de la cellule hôte (*E. coli*). Sous cette forme silencieuse, non-virulente, le virus est qualifié de prophage. L'intégration est réalisée à un locus spécifique du génome bactérien [Figure 8]. En somme, le phage intègre ou "clone" son génome dans un vecteur réplcatif (ici le chromosome entier) de *E. coli*. Le système Gateway™ [33] est une technique de clonage de laboratoire, dérivée de celle utilisée par le phage.

3.1.1.2. La technologie du MultiSite Gateway™

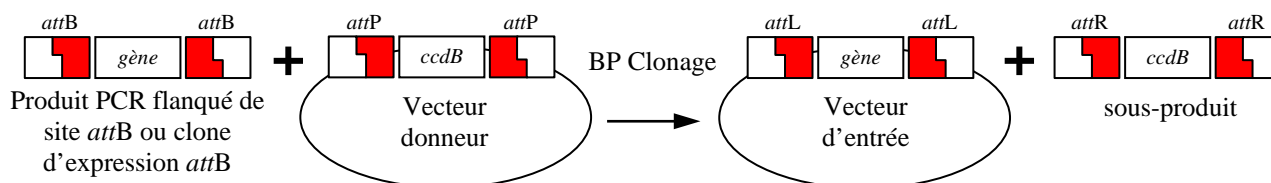
Le système Gateway™ est une technologie de clonage basée sur le système de recombinaison site-spécifique du bactériophage lambda qui fournit un moyen rapide et très efficace de transférer des séquences d'ADN hétérologues dans un système de vecteurs multiples en vue d'une analyse fonctionnelle ou de l'expression de protéines. Chez le bactériophage lambda, la recombinaison a lieu entre le phage et le chromosome de *E. coli* via des séquences de recombinaison spécifiques (sites *att* pour attachement), et est catalysée par plusieurs protéines de recombinaison. Les réactions sont décrites dans le tableau ci-dessous.

Voie	Réaction	Catalysée par...
Lysogénique	$attB \times attP \rightarrow attL \times attR$	Int, IHF (BP clonase™)
Lytique	$attL \times attR \rightarrow attB \times attP$	Int, Xis, IHF (LR clonase™)

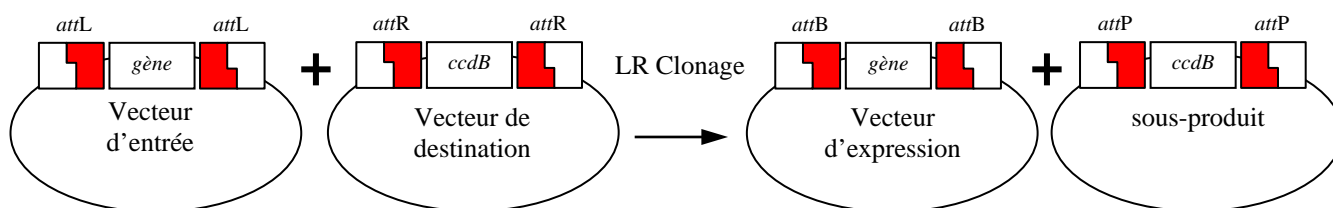
Une conséquence importante de la constitution des sites *att* est que les réactions d'intégration et d'excision ne mettent pas en jeu la même paire de séquences : ces réactions de recombinaison possèdent un caractère directionnel et, de plus, elles sont réversibles mais dans des conditions différentes.

La technologie Gateway™ utilise des sites *att* modifiés et optimisés pour permettre le transfert de séquences d'ADN hétérologues entre vecteurs. Deux réactions de recombinaisons constituent la base de la technologie Gateway™ :

- **Réaction BP** [Figure 9a] : elle facilite la réaction d'un substrat *attB* avec un substrat *attP* pour créer un clone d'entrée contenant des sites *attL*. Cette réaction est catalysée par un mix enzymatique BP Clonase™, qui est un mélange de l'intégrase du phage λ (Int) et de la protéine de *E. coli* Integration Host Factor (IHF).



- **Réaction LR** [Figure 9b] : elle facilite la recombinaison d'un clone d'entrée contenant des sites *attL* avec un vecteur de destination contenant des sites *attR* pour créer un clone d'expression contenant des sites *attB*. Cette réaction est catalysée par un mix enzymatique LR Clonase™, qui est un mélange de l'intégrase (Int) et de l'excisionase (Xis) du phage ainsi que de la protéine de *E. coli* Integration Host Factor (IHF).



Pour permettre le clonage recombinant avec le système Gateway™, les sites *att* sauvages du phage λ ont été modifiés pour améliorer l'efficacité et la spécificité des réactions de recombinaisons.

Dans le système MultiSite Gateway™, les sites *att* ont été optimisés pour permettre un clonage simultané de plusieurs fragments d'ADN. Ces modifications incluent des changements à la fois de séquences et de longueurs des sites *att* avec pour conséquence la création de nouveaux sites *att*. Ces sites ont une spécificité et une efficacité améliorées, chose qui est requise pour cloner plusieurs fragments d'ADN en même temps.

En général, les sites *att* modifiés de la technologie MultiSite Gateway™ montrent la même spécificité que ceux de la technologie Gateway™. C'est-à-dire que :

- Les sites *attB* réagissent seulement avec les sites *attP* ; par exemple, un site *attB3* réagit avec un site *attP3* pour générer un site *attL3*.
- Les sites *attL* réagissent uniquement avec les sites *attR* ; par exemple, un site *attL2* réagit avec un site *attR2* pour générer un site *attB2*.

Cependant, en fonction de l'orientation et de la position des sites *attB* et *attP* sur le fragment d'ADN d'intérêt et sur le vecteur donneur, respectivement, la réaction de BP clonage peut avoir pour conséquence la création d'un site *attR* au lieu d'un site *attL*. De manière spécifique, on observe que :

- Un site *attB1* réagit avec un site *attP1_R* pour générer un site *attR1*,
- Un site *attB2* réagit avec un site *attP2_R* pour générer un site *attR2*.

L'indice " R " signifie réverse. L'orientation (à partir du côté 5' vers le côté 3') de ce site est opposée à celle des sites *att* classiques.

Par exemple [Figure 10] , un produit PCR flanqué de sites *attB4* et *attB1* est utilisé pour faire une réaction de BP clonage avec un vecteur pDONR™ P4-P1R.



A cause de l'orientation et de la position des sites *attB1* et *attP1_R*, dans le produit PCR et dans le vecteur donneur, respectivement, le vecteur d'entrée résultant contient le produit PCR flanqué par un site *attL4* et un site *attR1* plutôt que deux sites *attL*.

Afin de permettre une sélection efficace des clones d'entrée ou d'expression, les vecteurs donneurs et de destination du système MultiSite Gateway™ contiennent deux sites *att* qui flanquent une cassette contenant :

- Le gène *ccdB* pour une contre-sélection (sélection sur l'évènement " perte de *ccdB* "),
- Le gène de résistance au chloramphénicol (*Cm^R*) pour une sélection positive.

Après une réaction de recombinaison MultiSite Gateway™, cette cassette est remplacée par le gène d'intérêt pour générer les clones d'entrées ou d'expression.

La présence du gène *ccdB* permet une sélection négative des vecteurs de destination ainsi que des vecteurs donneurs, dans *E. coli*, après recombinaison et transformation. La protéine *ccdB* interfère avec la DNA gyrase de *E. coli*, empêchant par là la croissance de la plupart des souches de *E. coli* (TOP10, DH5α™). Lors de la recombinaison, le gène *ccdB* est remplacé par le gène d'intérêt. Les cellules qui ont intégré un vecteur qui n'a pas réagi et qui donc contient le gène *ccdB* ou bien un sous-produit contenant le gène *ccdB* ne se multiplieront pas.

3.1.2. Acteurs et partenaires

3.1.2.1. La levure comme éprouvette d'étude

C'est la levure modèle *Saccharomyces cerevisiae* que nous utilisons comme outil pour la reconstruction *in vivo* des modules d'intérêt. La souche utilisée est la MaV203 construite par le Dr Marc Vidal [34, 35]. Cette souche est utilisée dans la technique du double-hybride en levure [Figure 11]. Elle contient :

- Des mutations d'auxotrophie : *leu2* et *trp1* pour permettre la sélection des deux plasmides exprimant les protéines hybrides, et *his3* pour la croissance dépendante de l'expression de *GAL1::HIS3* sur un milieu contenant du 3-AT.
- Des délétions des gènes *GAL4* et *GAL80* encodant Gal4p et son répresseur Gal80p, respectivement. En absence de Gal80p, le galactose n'est pas nécessaire pour l'activation des promoteurs inductibles par Gal4p.
- Trois gènes rapporteurs inductibles par Gal4p, en simple copie, intégrés de manière stable: *SPAL::URA3*, intégré au locus *URA3* ; *GAL1::HIS3*, intégré au locus *LYS2* ; *GAL1::lacZ*, intégré à un locus inconnu.

Des levures transformantes, exprimant des interactions protéines-protéines bien décrites, ont été caractérisées et sont utilisées comme contrôles phénotypiques dans les expériences de double-hybride. Il s'agit de :

- "contrôle 1", vecteur AD et DB vides, utilisé comme contrôle négatif.
- "contrôle 2", qui exprime une interaction relativement faible, DB-RB et AD-E2F.
- "contrôle 3", DB-Fos et AD-Jun, qui exprime une interaction relativement forte.
- "contrôle 4", Gal4p (longueur entière) et un vecteur AD vide.
- "contrôle 5", DB-DP et AD-E2F, qui exprime une interaction relativement forte.

Les différentes protéines sont placées sous le contrôle du promoteur de force modérée du gène de l'alcool déshydrogénase de levure (*ADH1*).

3.1.2.2. Les acteurs trans : acteurs protéiques

La régulation temporelle de l'expression des gènes au cours de la transition de phase G1 à S du cycle de division cellulaire est un processus compliqué dans lequel la formation d'un réseau d'interactions protéines/protéines impose un contrôle à plusieurs étapes. Les facteurs de transcription E2F et les "protéines-poches" de la famille du rétinoblastome RB (pRB, p107 et p130) sont des régulateurs clés du cycle de division cellulaire des eucaryotes supérieurs. E2F active la transcription de gènes impliqués dans la réplication de l'ADN et dans la progression du cycle de division. Quand les cellules sortent du cycle de division, E2F se lie aux "protéines-poches" qui, à leur tour, recrutent un nombre de corépresseurs (p.ex. le coactivateur CBP/P300) de la transcription aux promoteurs des gènes régulés par E2F [36]. En l'absence de "protéines-poches", les cellules ne sont plus capables d'arrêter leur cycle de division et deviennent immortelles. L'élimination de cette barrière à l'entrée du cycle de division cellulaire est l'une des étapes qui mènent à la formation de cancers.

Le premier indice vers la compréhension du rôle moléculaire des "protéines-poches" a été obtenu lorsqu'il a été remarqué que E1A (mais également E7 et SV 40 large T) une oncoprotéine virale d'un petit adénovirus, ciblait RB. [37] [38] [39] Cette oncoprotéine virale interagit avec le domaine poche de RB, p107 et de p130 par son motif LXCXE qui est essentiel pour ses fonctions de transformation [Figure 12]. Donc les oncogènes viraux interfèrent avec les partenaires normaux des "protéines-poches" et interfèrent vraisemblablement avec les fonctions d'arrêt du cycle cellulaire.

Le facteur de transcription E2F a été découvert à l'origine à cause de son activité favorisant la transcription du gène viral *E2*. Le gène *E2* est localisé dans la « **early 2** transcriptionnal region » du génome adénoviral et il encode une protéine de liaison à l'ADN requise pour la réplication de l'ADN adénoviral. La forme libre (non-liée à pRB) d'E2F ainsi que la transcription du gène *E2* sont favorisées par l'oncoprotéine E1A. Donc E1A semble inhiber la liaison de pRB à E2F et par là permettant à E2F libre d'activer la transcription d'*E2*.

La protéine RB suppresseur de tumeur joue un rôle central dans la régulation négative de la prolifération cellulaire et cette fonction d'inhibition de croissance dépend, au moins en partie, de sa capacité à réprimer l'activité de transcription d'E2F. Du début à la moitié de la phase G1, la protéine RB est phosphorylée et inactivée par des complexes de cyclines dépendants des kinases. Cet événement relâche E2F de RB, et E2F peut médier la prolifération cellulaire. L'importance de RB/E2F est soulignée entre autres par le fait que des oncoprotéines virales, telle E1A, inactivent les propriétés de suppresseur de tumeur de RB en interagissant avec RB de telle manière qu'E2F est libéré de RB.

3.1.2.3. Les éléments cis : éléments de séquence d'ADN

1°) Promoteur *Cup1*

L'utilisation de ce promoteur, " activé " en présence de cuivre (ON) et " inactivé " en absence de cuivre (OFF) permet d'activer de manière permanente la transcription d'ORF sous son contrôle. En plaçant le gène E1A sous le contrôle du promoteur *Cup1* et en présence de cuivre, la transcription d'E1A sera activée. Dans le contexte génétique du contrôle 2 (voir page précédente) utilisé dans la technique du double-hybride, la protéine E1A inhibe l'activité de DB-RB avec pour conséquence la disparition de l'activité β -Galactosidase (couleur bleue devient blanche).

Dans la perspective d'un réseau non-régulé (voir ci-dessous), l'idée sera de mettre le promoteur *Cup1* en position dite "ON " (présence de cuivre) et d'observer l'éventuel résidu d'activité β -Galactosidase qui constituera une mesure du bruit de fond.

Dans la perspective d'un réseau oscillant, l'idée sera de "synchroniser " toutes les cellules de levures dans leur cycle d'oscillation afin de minimiser les différences individuelles et d'obtenir une certaine homogénéité au niveau de toute une population. Une fois le cuivre retiré du milieu de culture, nous espérons observer une oscillation synchrone au niveau de l'entièreté de la population levurienne.

2°) Promoteur *Gall*

Dans la perspective d'un réseau oscillant, le but de l'utilisation du promoteur *Gall* est de mettre une ORF d'intérêt (E1A) sous le contrôle d'un facteur de transcription de telle manière que l'activité de celui-ci soit inhibée par la protéine encodée par l'ORF sous son propre contrôle (boucle de rétrocontrôle). Par ailleurs, il s'agit du même promoteur que le "readout " (*LacZ* pour un test β -Gal) ce qui permet de garder des conditions de transcription semblables entre ORF et "readout ".

3.2. Réseaux à reconstruire

Nous allons tenter de reconstruire, au départ des acteurs protéiques E1A/RB/E2F, deux réseaux présentant deux comportements différents. Notons que les protéines que nous utilisons ne correspondent pas aux versions sauvages. En effet, E2F (aa 342-437) et RB (aa 302-928) sont fusionnées respectivement au domaine AD (AD pour Activation Domain, aa 768-881) du facteur de transcription GAL4 et au domaine DB (DB pour DNA binding Domain, aa 1-147) du même facteur de transcription GAL4 [40]. Il sera, par la suite, toujours fait référence à AD-E2F et DB-RB. Par ailleurs, AD-E2F est également fusionné au SV40 NLS (NLS pour Nuclear Localization Sequence) qui permet à la protéine de pénétrer dans le noyau pour y réaliser son activité. La protéine de fusion DB-RB quant à elle comporte déjà un NLS dans le domaine DB [Figure 13].

3.2.1. Réseau non-régulé

On l'a dit, E1A inhibe l'activité de RB. Ce premier réseau envisagé voit le gène E1A placé sous le contrôle d'un promoteur inductible par le cuivre. En présence de cuivre, la transcription est ON, alors qu'en absence, elle est OFF. En position ON, en principe DB-RB ne sera jamais actif et on ne devrait pas avoir d'expression de la β -Galactosidase. Cependant, le contrôle peut avoir des "fuites " (" leakiness du contrôle ") [Figure 14]. Ce réseau permettra de mesurer le bruit de fond du système.

Ci-dessous, on trouvera la table des états des éléments de ce réseau.

E1A	DB-RB	AD-E2F	β -Galactosidase	
1	1	1	1	illégal
1	0	1	0	
0	1	1	1	
0	0	1	0	illégal

Dans cette table des états, on constate que :

- Deux des quatre états du réseau n'ont pas de "sens " et sont, pour cette raison, qualifiés "d'illégaux " C'est-à-dire que la combinaison des états des éléments ne correspond à aucune situation physiologique (E1A et DB-RB ne peuvent être actifs en même temps).
- Parmi les deux états restants, on voit que " **SI** E1A est 1 (entendons E1A est actif), **ALORS** DB-RB n'est **PAS** 1 (entendons DB-RB n'est pas actif) ".

3.2.2. Réseau oscillant

3.2.2.1. Présentation, Représentation booléenne et Table des états associés

Le gène E1A est placé sous le contrôle du promoteur *Gall*. Ce promoteur active la transcription des gènes sous son contrôle lorsque AD-E2F et DB-RB sont présents et actifs simultanément.

Ci-dessous, on trouvera la table des états des éléments de ce réseau.

E1A	DB-RB	AD-E2F	E1A
0	0	0	0
1	0	1	0
0	0	1	0
0	1	1	0
0	1	1	1
1	1	1	1
1	0	1	1
1	0	1	0
0	0	1	0
0	1	1	0

Les trois premières colonnes correspondent aux trois partenaires du réseau. La quatrième colonne correspond à " l'output " du réseau. Or, il s'avère que dans ce cas, l'output est également un des partenaires. Un tiret suscrit signifie un "ordre de commutation" c'est-à-dire que l'élément qui en est affecté verra la valeur de son état changée.

Chaque ligne du tableau correspond à une étape dans l'évolution de la dynamique du réseau (elles se suivent chronologiquement en partant du haut vers le bas). Considérons l'état hypothétique de départ pour lequel tous les partenaires ont la valeur zéro (inactif) et auxquels on donne simultanément un ordre de commutation. Sachant les relations qu'entretiennent les éléments les uns avec les autres, il est possible de prédire quel(s) élément(s) va voir son état modifié. La succession des lignes du haut en bas est une succession de changements intervenant dans le comportement du module à cause des interactions des éléments qui le composent. On se rend compte qu'au bout de quelques changements, on rencontre le même état du module (lignes grisées). Il existe une évolution cyclique du comportement du module.

Les relations logiques de ce réseau [Figure 15] peuvent être lues de la manière suivante :

- **Si** DB-RB **ET** AD-E2F sont 1 (entendons DB-RB et AD-E2F actifs simultanément), **alors** E1A est 1 (entendons E1A est actif).
- **Si** E1A est 1 (entendons E1A est actif), **alors** DB-RB n'est **PAS** 1 (entendons DB-RB n'est pas actif).

En connaissant les relations entre les éléments du réseau, il est possible d'en appréhender la dynamique : il s'agit d'une boucle, plus précisément d'une boucle de rétroaction négative.

3.2.2.2. *Simulation in silico*

La théorie des boucles de rétrocontrôle développée par R. Thomas nous dit que dans une boucle, si le nombre d'interactions négatives est impair, on parle de boucle négative ; un système à boucle négative tendra à osciller. Nous avons cherché à construire un réseau présentant un tel comportement. On peut effectivement constater sur la représentation booléenne [Figure 15] qu'il s'agit bien d'une boucle de régulation négative. Mais avant de tester *in vivo* le comportement du réseau, nous l'avons préalablement testé *in silico* [Figure 16]. On peut observer que le réseau présente un comportement oscillant. Si on transpose au comportement correspondant *in vivo*, on attend que, en présence de cuivre, E1A soit activé de manière permanente. Ceci permet de synchroniser les cellules dans un seul état du réseau. Si ensuite on enlève cet input, alors le réseau des cellules commencera à produire des oscillations ce qui devrait se traduire par un readout périodique de la population cellulaire.

*

* *

Biologiquement, l'étude de ce module de régulation simple choisi permet d'approcher deux problématiques : la première a trait à la régulation du cycle cellulaire, la seconde aux conséquences de l'activité d'E1A. Il est généralement admis que la régulation du cycle cellulaire est très complexe et fait intervenir un grand nombre d'éléments. De notre point de vue, il est intéressant de se demander quels sont les *éléments de régulation* nécessaire à l'explication de ce comportement cyclique. En d'autres mots, quelle est la logique de régulation qui sous-tend ce processus à la fois très constant et complexe. La disruption, par E1A, de l'interaction RB et E2F mime (et représente un énorme raccourci) l'action des phosphorylations inhibitrices sur RB par des CDK *in vivo*. Il sera intéressant plus tard de remplacer E1A par les acteurs réels du cycle cellulaire et de voir si le comportement prédit est observé. Autrement dit, une ultra-simplification pourrait permettre d'expliquer le comportement intégré et complexe d'un module. L'autre problématique abordée jette un regard sur l'action d'une protéine virale sur la dynamique d'un module. Il existe un parallèle entre mettre le gène E1A sous le contrôle du promoteur *Cup1* en présence de cuivre et l'activation d'E1A dans une cellule mammalienne infectée par un adénovirus. Dans les deux cas, E1A casse l'interaction de RB et E2F ce qui libère E2F.

Il a d'abord été nécessaire de construire les différents éléments de ce module pour pouvoir les assembler ensuite, en utilisant la technologie Gateway™ en un réseau, afin de reconstituer un module. La première étape a été la construction moléculaire d'un vecteur de destination qui soit compatible avec la technologie du MultiSite Gateway™ et qui puisse être sélectionné en levure. La seconde étape a été de "BP cloner" les éléments du module, promoteurs et ORFs, dans les vecteurs d'entrée du système Gateway™. Nous avons vu que par le LR Clonage, il est possible de les combiner à volonté. Il s'agit là d'une véritable boîte à outils dans laquelle nous pouvons puiser à souhait pour associer un promoteur et une ORF au choix et les placer en phase de lecture l'un avec l'autre.

1. PRÉPARATION DES ÉLÉMENTS DE CONSTRUCTION DU VECTEUR D'EXPRESSION D'E1A

1.1. Aménagement de la matrice

1.1.1. Vecteur de départ et Stratégie d'ensemble du clonage

1.1.1.1. Description du vecteur de départ pPC86

Le plasmide pPC86 [Figure 1] à partir duquel les constructions moléculaires ont été réalisées ne contient que certains des éléments dont nous avons besoin pour obtenir un vecteur d'expression fonctionnel en levure. Les caractéristiques de ce vecteur incluent :

- La séquence *ARS4* pour une réplication autonome dans la levure,
- La séquence centromérique *CEN6* pour la maintenance à un faible nombre de copies (low-copy-number)³,
- Le gène *TRP1* pour la sélection en levure sur un milieu déplété en tryptophane,
- Le promoteur constitutif pADH1, de force modérée, du gène de l'alcool déshydrogénase de levure (*ADH1*),
- Un site multiple de clonage (MCS),
- Le terminateur de transcription du gène de l'alcool déshydrogénase de levure (*ADH1*),
- Une origine de réplication (pUC ori) basée sur celle du plasmide pUC ainsi qu'un gène de résistance à l'ampicilline (Amp^R) pour la réplication et la maintenance dans *E. coli*.

Parmi les éléments bactériens que le plasmide porte, deux sont d'intérêt pour la construction du vecteur de destination : l'origine de réplication (pUC ori – ColE1) qui permet la transmission du plasmide d'une génération bactérienne à une autre, ainsi que le gène de résistance à l'ampicilline (avec son promoteur) qui permet une sélection (sur un milieu approprié) des bactéries qui possèdent ce plasmide. Les deux éléments levuriens intéressants sont d'une part l'origine de réplication levurienne (*CEN6/ARSH4*) qui permet la transmission du plasmide d'une génération de levure à une autre et, d'autre part, le terminateur de transcription situé juste en aval du site multiple de clonage.

³ Il semble que le faible nombre de copies soit une condition à la reconstitution de modules *in vivo*.

1.1.1.2. Stratégie

1°) Etape 1

La souche de levure (*S. cerevisiae*) utilisée (MaV 203) est déjà transformée par deux plasmides qui portent DB-RB et AD-E2F [34, 35]. Ces deux plasmides portent des gènes conférant la prototrophie pour la leucine et le tryptophane respectivement (deux marqueurs de sélection fréquemment utilisés en levure). Pour maintenir la sélection sur ces deux plasmides, les levures sont cultivées sur des milieux sans leucine et sans tryptophane de sorte que ne poussent que celles qui possèdent les deux plasmides. Dans cette souche, on trouve également trois gènes rapporteurs (*URA3*, *HIS3* et *LacZ*), en simple copie, intégrés de manière stable dans le génome. Ces marqueurs sont activés par GAL4 et peuvent être utilisés comme marqueurs de sélection. La sélection sur AD-E2F est maintenue grâce au marqueur de prototrophie *TRP1* (croissance en absence de tryptophane). Ce même marqueur est également présent sur pPC86. Puisqu'il s'agit de la même sélection, il est possible que les levures perdent un des deux vecteurs sans phénotype apparent (absence de croissance). Pour cette raison, nous avons excisé de pPC86 le gène (ainsi que son promoteur) *TRP1*.

2°) Etape 2

Le but de la seconde étape, de construction du vecteur de destination, est de le rendre compatible avec la technologie du MultiSite Gateway™. Ceci a été réalisé par intégration au site multiple de clonage d'une " cassette Gateway™ ". Cette cassette doit permettre une réaction de "LR clonage" entre le vecteur de destination construit, un vecteur d'entrée portant un promoteur d'intérêt et un vecteur d'entrée portant une ORF d'intérêt pour mettre le promoteur et l'ORF d'intérêts dans le vecteur de destination.

3°) Etape 3

Afin de garder l'utilisation de chacun des trois rapporteurs qui sont des marqueurs de sélection classique, un autre type de marqueur de sélection a été recherché. Il nous fallait un marqueur de sélection additionnel différent de ceux-ci. Un gène (associé à un promoteur bactérien, un promoteur levurien et un terminateur de transcription) conférant la résistance à la zéocine (*Sh ble*) a été cloné dans le vecteur déjà rendu compatible avec la technologie MultiSite Gateway™. Ceci va permettre la sélection, en levure, des cellules transformées par le vecteur de destination recombiné (vecteur d'expression).

La zéocine est un antibiotique de la famille de la bléomycine/phléomycine. Sa forte toxicité envers les bactéries, les champignons (dont la levure) et les cellules mammaliennes en fait un bon marqueur de sélection. La zéocine est un dérivé de la phléomycine D1, un glycopeptide basique, hydrosoluble et chélateur de cuivre, isolée de *Streptomyces verticillus*. La forme " cuivre-chélaté " est inactive. Lorsque l'antibiotique entre dans la cellule, le cation cuivre est réduit de Cu^{2+} en Cu^+ qui est retiré par des composés sulfhydryls de la cellule. La zéocine alors activée, se lie à l'ADN et le clive, provoquant la mort cellulaire. Une protéine de résistance à la zéocine a été isolée et caractérisée. Cette protéine de 13,665 kDa, produit du gène *Sh ble* (*Streptoalloteichus hindustanus*), se lie de manière stoéchiométrique à la zéocine et inhibe son activité de clivage de l'ADN. L'expression de cette protéine chez des hôtes procaryotes comme eucaryotes confère la résistance à la zéocine.

1.1.2. Etape 1 : Elimination du gène et du promoteur *TRP1*

La stratégie adoptée est une double restriction, de part et d'autre du fragment promoteur + gène *TRP1* (positions 3252 à 4029 sur pPC86), puis une "ligation" sur lui-même du vecteur dépourvu du fragment excisé. Les deux enzymes choisies sont BamHI et SnaBI.

L'endonucléase BamHI clive le vecteur pPC86 une seule fois en un site (position 2504) situé en aval (donc côté 3') du gène *TRP1*. Comme cette enzyme laisse des bouts collants et que SnaBI, l'autre enzyme choisie, laisse des bouts francs après coupure, nous avons du "remplir" les quelques nucléotides manquants (fill in). Une purification du produit de restriction a été réalisée (via un kit Qiagen PCR Purification Kit) afin d'éliminer les protéines (BamHI) et de permettre un changement de tampon.

Après cette purification, un *fill in* des bouts protrudants est effectué grâce au fragment Klenow de l'ADN Polymérase I. La taille du vecteur restreint est vérifiée sur un gel d'agarose en présence de bromure d'éthidium.

L'endonucléase SnaBI coupe le vecteur pPC86 une seule fois en un site (position 4072) en amont (côté 5') du promoteur de *TRP1*. Cette double restriction excise également *f1 ori* qui n'est pas utile (origine de réplication simple brin du phage M13). A cette étape également, le nombre ainsi que la taille des fragments du vecteur doublement restreint sont visualisés sur un gel d'agarose en présence de bromure d'éthidium.

Une fois ces deux restrictions effectuées, il reste à refermer le vecteur sur lui-même. Pour éviter que l'insert ne soit "ligué" à son site de départ, mais également pour éviter de récupérer les vecteurs non-digérés ou bien partiellement digérés, une extraction et purification du produit de double digestion à partir d'un gel d'agarose a été réalisée (via un kit QIAquick Gel Extraction Kit). Après migration et séparation les fragments d'ADN dans le gel, la bande d'intérêt est excisée puis l'ADN est purifié.

La réaction de "ligation" a ensuite été effectuée au moyen de la T4 DNA Ligase. Cinq μ l du mix de "ligation" sont utilisés pour transformer 100 μ l de bactéries thermocompétentes (souche DH5 α). Après transformation, dix clones indépendants sont prélevés au hasard pour commencer une Miniprep (inoculation de milieu LB liquide + ampicilline –50 μ g/ml– sur la nuit) (via un kit Qiaprep Spin Miniprep Kit). Nous avons appelé ce vecteur **pPC86MT1** (pour **pPC86 Minus Trp1**). Le contrôle de la taille du vecteur amplifié par Miniprep est basé sur une restriction par l'endonucléase NotI. Ce contrôle sert également de point de départ de l'étape suivante.

1.1.3. Etape 2 : Clonage de la cassette MultiSite Gateway™ dans le vecteur intermédiaire pPC86MT1

1.1.3.1. Cassette MultiSite Gateway™

La cassette bordée de sites *att* adéquats a été clonée dans pPC86MT1. Cette cassette est constituée, de 5' vers 3', d'un site *attR4*, du gène *Cm^R* (résistance au chloramphénicol), du gène *ccdB* et d'un site *attR2*.

Cette cassette nous a été aimablement donnée par le Dr Mike Boxem. Elle provient d'un plasmide de destination (pDEST6) qu'il a construit. Cette cassette a été excisée par restriction via l'endonucléase EcoRV. Cette enzyme laisse des bout francs après coupure.

1.1.3.2. Ouverture de pPC86MT1

Le vecteur pPC86MT1 a été restreint par l'endonuclease NotI qui coupe en un site unique (position 2033) situé juste en amont du terminateur de transcription (*ADHTT*). L'enzyme NotI laisse des bouts protrudants. Par ailleurs, la " cassette MultiSite Gateway™ " est, elle, coupée à bouts francs. Après purification du produit de restriction (via un kit Qiagen PCR Purification Kit) les quelques nucléotides manquant ont été remplis (*fill in*) grâce au fragment Klenow de la DNA Polymérase I pour permettre le clonage de la cassette qui est terminée par des bouts francs.

1.1.3.3. Déphosphorylation des extrémités 5'-phosphate

Pour éviter une fermeture du vecteur sur lui-même lors de l'étape de "ligation", il a été nécessaire de déphosphoryler les extrémités 5' phosphate du vecteur pPC86MT1 restreint par NotI. Cette étape est effectuée par la CIP alcaline phosphatase.

Notons qu'après chaque étape, le produit est purifié sur une colonne pour éliminer les protéines et permettre un changement de tampon d'activité (via un kit Qiagen PCR Purification Kit).

1.1.3.4. "Ligation" et transformation

La cassette "MultiSite Gateway™" est bordée de nucléotides phosphorylés en 5'. Après les étapes de restriction, *fill in*, et déphosphorylation du vecteur, nous avons "ligué" celui-ci avec la cassette MultiSite Gateway™ (dans un rapport 1 molécule de vecteur contre 10 molécules d'insert) via la T4 DNA Ligase. Après "ligation", 5 µl du mix sont utilisés pour transformer 100 µl de bactéries DB3.1 thermocompétentes. Les bactéries sont étalées sur un milieu LB supplémenté en ampicilline (50 µg/ml - sélection sur le vecteur) et en chloramphénicol (50µg/ml - sélection sur l'insert). Le vecteur ainsi obtenu est appelé pACR42 (pour **p**lasmi**d**e **A**RS4/**C**ENH6 **C**assette **R**4 **R**2). Après transformation, 6 clones indépendants sont prélevés au hasard et utilisés pour 6 Minipreps (inoculation de milieu LB liquide + ampicilline – 50 µg/ml et chloramphénicol – 50 µg/ml).

1.1.3.5. Contrôles

Trois contrôles ont été réalisés pour vérifier :

- La taille du vecteur,
- L'orientation de la cassette MultiSite Gateway™,
- La séquence de la cassette MultiSite Gateway™.

1°) Taille du vecteur et Orientation de la cassette MultiSite Gateway™

L'ADN plasmidique extrait des 6 clones a été digéré par deux enzymes de restriction qui permettent à la fois de déterminer dans quel sens est orientée la cassette MultiSite Gateway™ (parce qu'un clonage blunt a été réalisé, l'insert a pu être "ligué" dans l'un ou l'autre sens), et de déterminer si le profil de restriction donne des bandes en nombre et aux tailles attendues par la construction pACR42. La première enzyme, NotI, coupe à l'extrémité 3' de la cassette MultiSite Gateway™ du vecteur pACR42 (proche du site *attR4*). La seconde enzyme, KpnI, coupe quelques paires de bases en amont de la cassette. Si la cassette est orientée comme souhaité, c'est-à-dire R4 vers R2 de 5' vers 3' on s'attend à observer comme

profil de restriction sur un gel d'agarose, une bande 5350 bp et une bande à 2029 bp. Par contre, si la cassette est dans l'autre orientation, c'est-à-dire R2 vers R4 de 5' vers 3', on observera deux bandes de tailles plus ou moins équivalentes (3768 bp et 3611 bp). Seuls les clones répondant au premier profil de restriction seront utilisés pour la suite des constructions. Cinq des six clones prélevés au hasard présentent le bon profil.

2°) Séquence de la cassette MultiSite Gateway™

Afin d'être sûr de l'identité des sites *att* utilisés, nous avons décidé de les faire séquencer. Les primers utilisés pour le séquençage s'hybrident dans la cassette MultiSite Gateway™ et permettent un séquençage de l'intérieur vers l'extérieur de cette cassette. Les séquences reçues sont celles attendues.

1.1.4. Etape 3 : Clonage de la cassette zéocine pour la création de pZACR42

1.1.4.1. Amplification de la cassette zéocine par PCR

Nous avons à notre disposition un plasmide conférant la résistance à la zéocine (pPICZ◻A). Des primers (ZEO-FOR et ZEO-REV) s'hybrident aux extrémités de la cassette conférant la résistance à la zéocine ont été designés et ont permis une amplification par PCR de la cassette⁴ zéocine. Après réaction, quelques microlitres ont été déposés sur un gel d'agarose 1% (en présence de bromure d'éthidium) pour vérifier à la fois la présence et la taille de l'amplicon. Toutes deux sont correctes.

Nous avons essayé de cloner cet amplicon PCR dans le vecteur pACR42 restreint par deux enzymes (ApaI et AfeI) mais, à l'analyse des clones transformants, nous avons constaté qu'il ne s'agissait pas du vecteur de destination désiré. Pour cette raison, nous sommes passés par une étape supplémentaire pour cloner la cassette zéocine.

1.1.4.2. Clonage de la cassette zéocine dans un vecteur TOPO

L'idée était de cloner le produit de l'amplification PCR dans un vecteur conçu pour cloner des fragments d'ADN à bouts francs (vecteur pCR®- BluntII- TOPO®), d'ensuite l'en extraire par restriction pour enfin le cloner dans le vecteur pACR42 (via des bouts collants). La démarche de clonage de l'amplicon PCR peut être résumée en quatre étapes :

- Amplification de l'insert d'intérêt par PCR
- Réaction de TOPO® - Clonage
- Transformation dans des cellules compétentes
- Sélection et analyse des clones.

Le détail du fonctionnement de ce système de clonage est expliqué dans le chapitre consacré aux matériels et méthodes. La technologie du TOPO-Clonage se base sur la capacité d'une DNA Topoisomérase virale de transférer l'énergie d'un lien phosphate-sérine à un nucléotide et par là de permettre la création d'un lien entre une extrémité 3'-phosphorylée (le phosphate qui se trouve sur la sérine de l'enzyme) à un groupement hydroxyl d'une extrémité 5' nucléotidique. Pour la réaction de TOPO-clonage, il est nécessaire de mettre en présence produit PCR et vecteur TOPO linéarisé (fourni comme tel par *Invitrogen*). Notons que le produit PCR peut être "ligué" dans deux sens (voir plus loin dans le texte). Après la réaction

⁴ Par cassette zéocine, nous entendons la séquence nucléotidique qui est constituée de 5' vers 3' du promoteur levurien pTEF1, du promoteur bactérien pEM7, du gène *Sh ble* et du terminateur de transcription *CYC1* TT.

de clonage, qui ne prend que cinq minutes, 100 μ l de bactéries thermocompétentes et sensibles au produit du gène *ccdB* (TOP10) ont été transformées par deux microlitres du produit de "ligation". Les bactéries ont été étalées sur un milieu supplémenté en kanamycine (50 μ g/ml). Une double sélection est effectuée sur les bactéries:

- Sélection sur l'événement de transformation par la résistance à la kanamycine,
- Sélection sur l'événement d'insertion : la souche de bactéries utilisée est sensible au produit du gène *ccdB*. Si l'insert a été "ligué" correctement, c'est-à-dire entre le promoteur *Lac* et la protéine de fusion lacZ α -*ccdB*, il n'y a pas expression de cette dernière et les bactéries poussent. Dans le cas contraire, le produit du gène *ccdB* provoque la mort bactérienne.

Différents volumes de bactéries (10 μ l, 50 μ l et 240 μ l) sont étalés pour s'assurer qu'au moins une des boîtes présentera des colonies bien individualisées. Une transformation de la même souche de bactéries par le plasmide pCR®-Blunt II-TOPO® sans l'insert zéocine constitue le contrôle négatif de cette expérience.

Il est possible que la PCR ait introduit des mutations au cours des cycles d'amplification (bien qu'une polymérase à haute fidélité soit utilisée). Pour cette raison, sur la boîte présentant le plus grand nombre de colonies, douze sont choisies aléatoirement (pour augmenter les chances de prendre un clone sans mutation dans la cassette zéocine) et constituent le point de départ de deux étapes:

- Une Miniprep pour obtenir de grande quantité de plasmide
- Une PCR pour vérifier la présence de l'insert ainsi que sa taille (primers identiques à ceux utilisés pour amplifier la cassette zéocine).

Sur les douze colonies prélevées au hasard, quatre se sont révélées conformes pour la taille et la présence de la cassette zéocine.

1.1.4.3. Clonage de l'insert zéocine⁵ dans le vecteur pACR42 pour obtenir pZACR42

Afin d'augmenter l'efficacité de "ligation" par rapport aux clonages à bouts francs, un clonage par bouts collants a été réalisé, ce qui signifie qu'il était nécessaire de trouver des enzymes de restriction capables de couper correctement les deux plasmides à savoir pACR42 et pCR®-Blunt II-TOPO® + insert. Le vecteur pACR42 a donc été restreint en amont (donc côté 5') du promoteur *ADHI* par l'enzyme de restriction *ApaI* (position 94). Ce premier produit de restriction a ensuite été purifié (via un kit QIAgen PCR purification Kit). Ceci purifie l'ADN et permet de travailler pour chaque enzyme dans le tampon propre qui lui convient. Après purification, le vecteur coupé une fois a été restreint en aval du promoteur *ADHI* (côté 3') par l'endonucléase *HindIII* (position 1567). Le produit de la seconde digestion a lui aussi été purifié (via le QIAgen PCR purification Kit) pour ensuite voir ses extrémités 5' déphosphorylées par la CIP alcaline phosphatase⁶; ceci évite la "religation" du plasmide sur lui-même lors du clonage de la cassette zéocine.

Les produits de Miniprep ont été restreints (pCR® - Blunt II-TOPO® + insert clones 1 à 12) par l'enzyme *ApaI* qui coupe en aval de l'insert (côté 3' – position 276). Ensuite, et pour les raisons citées plus haut, le produit de la première digestion a été purifié (via le QIAgen PCR purification Kit). Le vecteur pCR® - Blunt II-TOPO® + insert coupé une fois, a finalement été restreint par l'endonucléase *HindIII* qui coupe en amont de l'insert (côté 5' - position 1566).

⁵ Insert zéocine = cassette zéocine sortie du vecteur pCR® – BluntII – TOPO® par restriction par *ApaI* et *HindIII*.

⁶ La déphosphorylation des extrémités 5'- phosphate prévient la fermeture du vecteur sur lui-même lors de l'étape de "ligation".

Le produit de la double digestion a également été purifié (via le QIAgen PCR purification Kit).

L'étape finale de la construction de ce plasmide de destination a été la "ligation" de la cassette zéocine avec le vecteur MultiSite Gateway™ compatible, pACR42. Le produit de digestion de pACR42 par *ApaI* et *HindIII* puis de déphosphorylation par la CIP alcaline phosphatase et les douze produits de digestion de pCR® - Blunt II-TOPO® + insert par *ApaI* et *HindIII* ont donc été mis en présence et ont subi la réaction de "ligation". De chaque "ligation", 5µl ont servi pour transformer des bactéries (DB3.1). Les produits de transformation sont étalés sur un milieu contenant zéocine (25µg/ml) et ampicilline (50µg/ml) ce qui permet une sélection sur les deux fragments mis en présence à l'étape de "ligation" (ampicilline pour le vecteur et zéocine pour l'insert).

1.1.4.4. *Tests de fonctionnalité en levure*

Le test de fonctionnalité que nous réalisons a pour but de vérifier que le plasmide que nous avons construit confère bien à la levure la résistance à la zéocine et deuxièmement, de déterminer ma fourchette de zéocine utilisable en sélection.

Pour cette raison, nous avons transformé la souche par le plasmide de destination construit (douze clones) et étalé ces levures sur des milieux sans leucine, sans tryptophane mais en présence de différentes concentrations de zéocine. Il apparaît qu'à une concentration de zéocine supérieure à 200 µg/ml même les levures transformées par le vecteur de destination ne poussent plus (concentration toxique). Les levures non-transformées, quant à elles, ne poussent plus à des concentrations de 50 µg/ml.

Deux contrôles supplémentaires ont été faits:

- Profil de restriction pour déterminer si le profil obtenu correspondait au profil attendu et vérifier l'orientation de la cassette zéocine
- Séquence de la cassette zéocine.

Sur onze des douze clones (un des douze clones a été perdu c'est-à-dire ne poussait pas lors de la Miniprep), une "restriction diagnostique" (diagnostic de taille et de nombre de fragments) a été effectuée. Les vecteurs de destination ont donc été restreints par l'endonucléase *NotI* qui coupe dans la cassette Gateway (une fois en dedans, une fois en dehors). Six clones sur les onze présentent le profil attendu, à savoir deux fragments de tailles correctes (donc orientation correctes de la cassette Gateway). Une seconde restriction diagnostic a été réalisée pour déterminer l'orientation de la cassette zéocine. Les onze clones ont été restreints par deux enzymes : *XbaI* et *EcoRV*. Seul les six clones déjà discriminés sur base de la première restriction présentent le profil attendu, à savoir quatre fragments de tailles correctes.

1.2. Création des mini-banques de promoteurs et d'ORFs

La technologie MultiSite Gateway™ nous permet d'assembler, au choix, un promoteur d'intérêt à une ORF d'intérêt. Nous avons voulu construire deux mini-banques, c'est-à-dire deux collections de vecteurs d'entrée qui contiennent d'une part tous les promoteurs d'intérêt (un mini-promoterome) et d'autre part toutes les ORFs d'intérêt (un mini-ORFéome). Le qualificatif de "mini" réfère au fait que par comparaison avec l'ORFéome (12000 ORFs clonées) et le Promoterome de l'organisme modèle *C.elegans*, le nombre d'éléments est très réduit. En effet, le mini-promoterome est constitué de deux vecteurs d'entrée portant le promoteur *Cup1* ou *Gall*, le mini-ORFéome quant à lui est

également constitué de deux vecteurs d'entrée qui portent soit le gène E1A c'est-à-dire CR1 et CR2 (version longue selon notre dénomination) soit E1A-CR2 (version courte selon notre dénomination).

1.2.1. Construction du mini-promoterome

1.2.1.1. Adaptation des promoteurs pour le BP clonage : amplification PCR

La première étape d'une réaction de BP clonage est une étape d'adaptation des promoteurs au système MultiSite Gateway™. Elle consiste en une amplification par PCR des promoteurs en utilisant des primers dont la partie 3' est spécifique de la séquence à amplifier et dont la partie 5' constitue un site de recombinaison *att*. Le site *att* du "primer forward" est un site *attB*₄, celui du "primer reverse" est un site *attB*_{1R}.

Nous avons décidé de considérer comme promoteur les 450 bp en amont de l'ATG du gène *Cup1*. Le promoteur *Cup1* a été amplifié à partir d'ADN génomique de levure (aimablement fourni par le laboratoire du Dr Pam Silver - DFCI).

Le promoteur *Gall* a été amplifié à partir du plasmide pPS293 (aimablement fourni par le laboratoire du Dr Pam Silver – DFCI). Nous avons considéré comme promoteur la séquence utilisée par la société *Invitrogen* dans ses vecteurs (450 bp). Quelques μ l du produit PCR sont déposés sur un gel d'agarose 1% pour vérifier la présence et la taille de l'amplicon.

1.2.1.2. Réaction de BP clonage proprement dite

Il s'agit de l'étape de recombinaison. Le produit PCR flanqué des sites *attB* adéquats est mis en présence d'un vecteur donneur, pDONR P4P1_R, ainsi que du mix BP Clonase™ (constitué de l'intégrase λ et de l'IHF de *E. coli*). La recombinaison permet "l'échange" entre le promoteur d'intérêt et les gènes *ccdB* et *Cm^R*.

L'étape suivante est la transformation de bactéries sensibles à *ccdB* (souche DH5 α) par deux μ l du mix de recombinaison. Les bactéries sont étalées sur un milieu contenant de la kanamycine (50 μ g/ml). Le vecteur donneur porte également un gène de résistance à la kanamycine. Ne poussent que les bactéries qui ont été transformées :

- par le vecteur donneur (sélection par la kanamycine)
- par le vecteur donneur recombiné (sélection par la kanamycine et par perte de la sensibilité de la souche à *ccdB*).

Une transformation par pDONR P4P1_R sans insert constitue le contrôle négatif de cette expérience. Le vecteur obtenu après recombinaison est appelé vecteur d'entrée. Nous avons donc pENTRY-*Cup1* et pENTRY-*Gall*. Les promoteurs sont bordés en 5' d'un site *attL*₄ et en 3' d'un site *attR*_{1R}. Quatre clones indépendants sont prélevés au hasard et constituent le point de départ de quatre Miniprep. Afin de vérifier la présence et la taille de l'insert, une PCR est effectuée en utilisant des primers qui s'hybrident en amont du site *attL*₄ d'une part (pDONR201-FOR) et en aval du site *attR*_{1R} (pDONR201-REV). On s'attend à voir pour les deux promoteurs une bande à environ 750-800 bp sur un gel d'agarose 1%, ce qui est effectivement observé.

De plus, ces deux vecteurs d'entrée ont vu la région comprise entre les deux sites *att* séquencée pour confirmer leur identité. Les séquences obtenues sont telles qu'attendues.

1.2.2. Construction du mini-ORFéome

1.2.2.1. Choix des ORFs

1°) E1A version longue (E1A FL)

Cette version de la protéine E1A ne correspond pas à la version WT. La protéine adénoviral (AD-5) E1A 12S wild-type est une protéine de 243 acides aminés. Notre version ne comprend que les deux domaines conservés (**CR** pour **C**onserved **R**egion) CR1 (aa 1 à 85) et CR2 (aa 121 à 139) nécessaires et suffisants à la transformation de cellules mammaliennes ainsi qu'à l'activité de disruption de l'interaction RB/E2F. A cette protéine, se trouve également ajouté en fusion un NLS.

2°) E1A version courte (E1A CR2)

Cette version courte ne comprend que CR2 (aa 76 à 139). Cette version ne doit avoir aucune activité de disruption de l'interaction RB/E2F. elle possède toutefois une très faible activité résiduelle. Il s'agit donc là d'un contrôle négatif. A cette protéine, se trouve également ajouté en fusion un NLS.

1.1.2.2. Réaction de BP Clonage

1°) Adaptation des ORF's pour le BP Clonage

On l'a dit plus haut, la première étape d'une réaction de BP clonage est une étape d'adaptation au système MultiSite Gateway™. Elle consiste en une amplification par PCR en utilisant des primers dont la partie 3' est spécifique de la séquence à amplifier et dont la partie 5' constitue un site de recombinaison *att*. Le site *att* du primer forward est un site *attB1*, celui du primer reverse est un site *attB2*.

Les deux versions de E1A utilisées ont été amplifiées à partir de deux plasmides construits par le Dr Ali Fattaey .

Quelques µl du produit PCR sont déposés sur un gel d'agarose 1% pour vérifier la présence et la taille de l'amplicon.

2°) Réaction de BP Clonage proprement dite

Il s'agit de l'étape de recombinaison. Le produit PCR flanqué des sites *attB* adéquats est mis en présence d'un vecteur donneur, pDONR 201, ainsi que du mix BP Clonase™ (constitué de l'intégrase \square et de l'IHF de *E. coli*). La recombinaison permet " l'échange " entre l'ORF d'intérêt et les gènes *ccdB* et *Cm^R*. La recombinaison se passe sur la nuit.

L'étape suivante est la transformation de bactéries sensibles à *ccdB* (souche DH5α) par deux \square l du mix de recombinaison. Les bactéries sont étalées sur un milieu contenant de la kanamycine (50 µg/ml). Le vecteur donneur porte également un gène de résistance à la kanamycine. Une double sélection est effectuée :

- Sélection de l'évènement transformation par la résistance à la kanamycine,
- Sélection de l'évènement échange de la cassette Gateway (et donc du gène *ccdB*) par la sensibilité de la souche bactérienne au produit du gène *ccdB*.

Une transformation par pDONR 201 sans insert constitue le contrôle négatif de cette expérience. Le vecteur obtenu après recombinaison est appelé vecteur d'entrée. Nous avons

donc pENTRY-*E1A-FL* (FL pour Full Length – notre version longue) et pENTRY-*E1A-CR2*. Les ORFs sont bordés en 5' d'un site *attL1* et en 3' d'un site *attL2*. Quatre clones indépendants sont prélevés au hasard et constituent le point de départ de quatre Miniprep's. Afin de vérifier la présence et la taille de l'insert, une PCR est effectuée en utilisant des primers qui s'hybrident en amont du site *attL1* d'une part (pDONR201-FOR) et en aval du site *attL2* (pDONR201-REV). On s'attend à voir pour les deux ORFs une bande à environ 750-800 bp sur un gel d'agarose 1%, ce qui est effectivement observé. De plus, ces deux vecteurs d'entrée ont vu la région comprise entre les deux sites *att* séquencée pour confirmer leur identité. Les séquences obtenues sont telles qu'attendues.

2. ASSEMBLAGE DES PIÈCES DU VECTEUR D'EXPRESSION D'E1A

Il s'agit de la réaction de MultiSite Gateway™ proprement dite. Avant de transformer des bactéries par le mix de recombinaison, l'expérience montre qu'il est préférable de restreindre le vecteur de destination avec une endonucléase qui coupe uniquement à l'intérieur de la cassette MultiSite Gateway (permet une recombinaison plus efficace et diminution du bruit de fond). L'enzyme choisie est BspEI qui coupe dans le gène de résistance au chloramphénicol (*Cm^R*).

Deux vecteurs d'entrée, l'un correspondant au vecteur entrée-promoterome l'autre au vecteur entrée-ORFéome, recombinent avec le vecteur de destination restreint que nous avons construit (pZACR42) grâce au mix enzymatique LR Clonase™ (constitué de l'intégrase λ , de l'excisionase Xis et de l'IHF de *E. coli*). La recombinaison permet " l'échange " entre la cassette *ccdB* et *Cm^R* avec le couple promoteur/ORF d'intérêts. La recombinaison se passe sur la nuit.

Quatre vecteur d'expression vont être construits dans lesquels on trouve quatre combinaisons de promoteur et d'ORF d'intérêts:

- E1A-FL sous le contrôle du promoteur *Cup1*,
- E1A-CR2 sous le contrôle du promoteur *Cup1*,
- E1A-FL sous le contrôle du promoteur *Gall* et,
- E1A-CR2 sous le contrôle du promoteur *Gall*.

L'étape suivante est la transformation de bactéries sensibles à *ccdB* (souche DH5 α) par deux \square 1 du mix de recombinaison. Les bactéries sont étalées sur un milieu contenant de l'ampicilline (50 μ g/ml). Le vecteur de destination porte également un gène de résistance à l'ampicilline. Une double sélection est effectuée :

- Sélection de l'évènement transformation par la résistance à l'ampicilline,
- Sélection de l'évènement échange de la cassette gateway (et donc du gène *ccdB*) par la sensibilité de la souche bactérienne au produit du gène *ccdB*.

A l'occasion du LR clonage, nous avons effectué plusieurs contrôles. Une transformation de bactéries DH5 α \square par le vecteur de destination coupé par BspEI constitue le contrôle négatif. D'autre part, afin de montrer la différence entre souche sensible et non-sensible à *ccdB*, nous avons également transformé des bactéries DB3.1 par le même mélange de restriction du vecteur de destination. Dans le même ordre d'idées, nous avons transformé des bactéries DH5 α ainsi que des bactéries DB3.1 par le vecteur de destination non-restreint. Ceci montre que :

- La souche DH5 α est bien sensible à *ccdB*, alors que la souche DB3.1 ne l'est pas
- Il existe « un bruit de fond » dans la souche DH5 α , c'est-à-dire que malgré la sensibilité, un certain nombre de colonies poussent.

Le vecteur obtenu après recombinaison est appelé vecteur d'expression (ou clone d'expression).

Quatre clones indépendants sont prélevés au hasard et constituent le point de départ de quatre Miniprep's. Afin de vérifier la présence et la taille de l'insert, une PCR est effectuée en utilisant des primers qui s'hybrident en amont du site *attB4* d'une part (primer forward = Gal4AD) et en aval du site *attB2* (primer reverse = ADHTT). Sur gel d'agarose 1%, les tailles des bandes observées sont telles que prédites.

3. RECONSTRUCTION ET TEST FONCTIONNEL *IN VIVO*

3.1. Choix du "readout"

Dans la souche de levure utilisée, on trouve les trois gènes rapporteurs: *HIS3*, *URA3* et *LacZ* par lesquels une sélection ou une détection peut être effectuée. Les deux premiers constituent des tests de croissance sur milieu sélectif, le troisième est un test visuel de détection de la couleur bleue. Les cinq contrôles du double-hybride servent de contrôle pour les expériences (phénotypes connus et bien décrits). Le contrôle 1 a une valeur de contrôle positif, c'est-à-dire où le rapporteur est OFF comme il devrait l'être si E1A joue son rôle d'interférence dans le réseau, puisqu'il n'y a pas de recombinaison du facteur de transcription *GAL4* (il s'agit du "bruit de fond" de la souche de levures). Le contrôle 2, qui contient DB-RB et AD-E2F, constitue le contrôle négatif c'est-à-dire qu'il correspond à la situation où E1A n'a pas d'effet. On sait qu'il n'y a pas inhibition de l'interaction DB-RB et AD-E2F.

3.2. Réseau non-régulé

3.2.1. Reconstruction dans la levure

La reconstruction dans la levure a été réalisée grâce à la transformation de la souche de levure MaV203 – contrôle 2 par les différents vecteurs d'expression générés par la technologie du MultiSite Gateway™ lors de la réaction de LR Clonage [voir Figure 10]. Après transformation, les levures ont été étalées sur un milieu sélectif (–Leu, –Trp, + Zéocine – 150 μ g/ml). C'est à partir de quelques clones transformés que les différents tests phénotypiques ont été réalisés.

3.2.2. Tests phénotypiques : révélation de la transactivation

Dans la souche transformée par *Cup1*-E1A (FL ou CR2), et en absence de cuivre, les trois gènes rapporteurs sont transcrits. Le contrôle 2 possède donc un triple phénotype :

- Sensibilité au 5-FOA (à cause de la transcription de *URA3*),
- Prototrophie pour l'histidine sur un milieu –His +3AT (à cause de la transcription de *HIS3*),
- Couleur bleue en test β -Gal (à cause de la transcription de *LacZ*).

Dans le réseau non-régulé, le gène E1A est placé sous le contrôle du promoteur *Cup1*. En présence de cuivre, E1A est toujours transcrit et traduit avec pour conséquence attendue une inhibition permanente de l'activité de DB-RB. Il ne devrait donc pas y avoir de recombinaison

du facteur de transcription *GAL4*. La question est de savoir si E1A-FL inhibe effectivement l'activité de DB-RB dans le réseau reconstruit. Il est possible d'observer les conséquences de l'activité d'E1AFL au plan phénotypique. On s'attend en effet à ce que les levures transformées présentent:

- Une résistance au 5-FOA,
- Une sensibilité au 3AT sur un milieu –His +3AT,
- Une couleur blanche en test β -Gal.

Il a par ailleurs été nécessaire de déterminer quelle concentration en cuivre était optimale pour activer la transcription via le promoteur *Cup1*. Dans une étude préalable, des concentrations de l'ordre de 0.25 mM sont montrées comme étant suffisantes pour activer la transcription via *Cup1*. Dans le but de déterminer quelle concentration en cuivre était optimale pour notre souche, nous avons réalisé les trois tests sur des milieux à concentration en cuivre variable (gamme de concentration de 0.0 mM à 0.75 mM en nitrate de cuivre). Ci-dessous, un tableau présente les différents milieux correspondant aux différentes conditions de test.

	–Leu –Trp	YEPD	–Leu –Trp + 5-FOA	–Leu –Trp –His + 3AT
zéocine 150 µg/ml	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre
sans zéocine	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre

Les levures transformées ainsi que les contrôles sont mises sur un milieu sans leucine et sans tryptophane (–Leu –Trp), en présence ou en absence de zéocine, à 5 doses différentes de cuivre. Elles sont ensuite mises à pousser 24 heures à 30°C. Différentes répliques [Figure] de ces boîtes-matrices sont réalisées sur les trois milieux qui permettent de faire les tests. Dans l'ordre, il s'agit d'une "réplique-transfert" (replica-plating) sur YEPD, puis deux répliques-épuisement (replica-cleaning) et une réplique-transfert sur un milieu –Leu –Trp –His + 3AT et enfin une réplique-transfert sur un milieu –Leu –Trp +5-FOA.

3.2.2.1. Tests de croissance

Test de croissance sur un milieu –Leu –Trp–His + 3AT

Sur le milieu qui est supplémenté en zéocine, et à la concentration maximale de cuivre, on observe une diminution de la croissance des levures transformées par *Cup1*-E1AFL par rapport aux levures transformées par *Cup1*-E1ACR2. Par ailleurs, sur un milieu supplémenté en zéocine mais en absence de cuivre, les levures transformées par *Cup1*-E1AFL ne présentent pas de différence de croissance par rapport aux levures transformées par *Cup1*-E1ACR2.

Sur un milieu sans zéocine et à concentration maximale de cuivre, aucune différence n'est observée entre les levures transformées par *Cup1*-E1AFL par rapport aux levures transformées par *Cup1*-E1ACR2. De même, sur un milieu sans cuivre, aucune différence de croissance n'est observée entre les levures transformées par *Cup1*-E1AFL et les levures transformées par *Cup1*-E1ACR2.

C'est seulement sur un milieu supplémenté en zéocine et à la concentration la plus élevée en cuivre qu'il est possible de voir un effet de E1AFL sur la croissance des levures. On peut dire que le cuivre a activé la transcription de E1A via le promoteur *Cup1* avec pour effet que E1AFL a "disrupté" (en partie) l'interaction de DB-RB et AD-E2F ce qui provoque une perte d'expression du gène rapporteur *HIS3*. On observe une diminution de croissance plus marquée par rapport avec E1AFL que dans le cas de levures transformées par *Cup1*-E1ACR2. La forme tronquée ne présente que très faiblement cette activité de "disruption", elle sert de contrôle positif interne.

Test de croissance sur un milieu –Leu –Trp+5-FOA

Sur un milieu supplémenté en zéocine et à la concentration maximale en cuivre, une faible croissance des levures transformées par *Cup1*-E1AFL par rapport aux levures transformées par *Cup1*-E1ACR2 est observée. D'autre part, sur un milieu supplémenté en zéocine mais en absence de cuivre, aucune différence de croissance entre ces deux souches n'est observable (absence d'activation d'E1A).

Sur un milieu sans zéocine et à concentration maximale de cuivre, une légère croissance des levures transformées par *Cup1*-E1AFL est visible par rapport aux levures transformées par *Cup1*-E1ACR2. Par contre, sur ce même milieu mais en absence de cuivre, aucune différence entre ces deux souches de levures n'est observée.

Dans ce test, on observe sur un milieu à la plus forte concentration en cuivre, supplémenté ou non en zéocine, une croissance des levures transformées par *Cup1*-E1AFL alors qu'aucune croissance n'est observable pour les levures transformées par *Cup1*-E1ACR2. On peut dire, ici encore, que le cuivre a activé la transcription de E1A via le promoteur *Cup1*. E1AFL a donc "disrupté" (en partie) l'interaction de DB-RB et AD-E2F ce qui provoque une perte d'expression du gène rapporteur *URA3* et donc une croissance par rapport aux levures transformées par *Cup1*-E1ACR2.

3.2.2.2. Test β -Gal

Sur un milieu supplémenté en zéocine, les résultats ne donnent aucune information valable. En effet, il est impossible de voir une différence significative entre les colorations. Par contre, sur un milieu sans zéocine, et à la concentration maximale en cuivre, une différence de coloration des levures transformées par *Cup1*-E1AFL par rapport aux levures transformées par *Cup1*-E1ACR2 est observée clairement au moins sur le document original couleur. Par contre, sur un milieu sans cuivre, aucune différence entre les deux souches n'est observable.

Sur un milieu sans zéocine, c'est uniquement à la concentration la plus élevée en cuivre qu'il est possible de voir un effet de E1AFL sur la croissance des levures. On peut dire à nouveau que le cuivre a activé la transcription d'E1A via le promoteur *Cup1*. E1AFL a donc "disrupté" (en partie) l'interaction de DB-RB et AD-E2F ce qui provoque une perte d'expression du gène rapporteur *LacZ* et donc une diminution de la coloration bleue par rapport aux levures transformées par *Cup1*-E1ACR2.

3.2.3. Cinétique de mise en place des acteurs

Pour déterminer après combien de temps la lecture du test β -Gal doit être réalisée, c'est-à-dire après combien de temps les différents éléments du réseau sont présents et exercent leur influence l'un sur l'autre, une "cinétique" du "readout" a été réalisée. Le test est effectué sur des levures transformées par le vecteur d'expression qui porte *Cup1-E1A* (FL ou CR2). Les levures sont déposées sur un milieu YEPD solide additionné de cuivre (0.25 mM) puis sont testées pour l'expression de la β -Galactosidase (test β -Gal). Les levures sont laissées à pousser sur milieu YEPD à 30°C pendant différents temps (en heures): 0, 2, 5, 19, 29 et 48. Aux trois premiers temps, aucune coloration des levures transformées par les constructions n'est visible. Par contre, pour les périodes les plus longues, on observe une décoloration des levures transformées par le vecteur d'expression *Cup1-E1AFL* par rapport aux levures non-transformées. Les levures transformées par *Cup1-E1ACR2* présentent également une légère décoloration.

Une mesure de la DO (Densité Optique) a été réalisée via le logiciel *Quantity one*® sur les images digitalisées des boîtes. Le contrôle 1 est considéré comme le contrôle positif de l'expérience : il n'y a en effet pas de reconstitution du facteur de transcription *GAL4*. La valeur de 100 % est attribuée au contrôle 2 non-transformé (contrôle négatif). La diminution observée par rapport à ce contrôle 100 % dans le cas des levures transformées par le vecteur d'expression qui porte *Cup1-E1AFL* est de 40 % ; dans le cas des levures transformées par *Cup1-E1ACR2*, elle est de 20 % du contrôle.

Un contrôle supplémentaire est réalisé. Il s'agit des mêmes levures (duplicat) cultivées sur milieu YEPD solide mais en absence de cuivre. Dans ce cas-ci, aucune différence n'est observée entre les levures transformées par *Cup1-E1AFL* ou *Cup1-E1ACR2* et les cinq contrôles du double-hybride.

1. Construction du vecteur d'expression d'E1A

Les caractéristiques du vecteur de destination, pZACR42 incluent:

- Une origine de réplication (pUC ori) basée sur celle du plasmide pUC ainsi qu'un gène de résistance à l'ampicilline (Amp^R) pour la réplication et la maintenance dans *E. coli*.
- Un gène de résistance à la zéocine pour la sélection en levure.
- Une cassette MultiSite GatewayTM qui permet la recombinaison par ce système.
- Le terminateur de transcription du gène de l'alcool déshydrogénase.
- La séquence *ARS4* pour une réplication autonome en levure.
- La séquence centromérique *CEN6* pour la maintenir à un faible nombre de copies.

Fonctionnellement:

- Confère la résistance à la zéocine aux bactéries transformées
- Confère la résistance à la zéocine aux levures transformées
- Permet la recombinaison dans la technologie MultiSite GatewayTM

Dans le réseau oscillant, E1A est placé sous le contrôle de deux promoteurs : *Gall* et *Cup1* (sur deux vecteurs différents). La construction *Gall*-E1A a été réalisée. Elle est portée par le vecteur pZACR42 recombiné (c'est-à-dire après réaction de LR Clonage). La sélection sur ce vecteur d'expression est effectuée via la zéocine. Pour reconstruire complètement le réseau oscillant, il est nécessaire de construire un second vecteur de destination qui puisse porter la construction *Cup1*-E1A. Il est donc nécessaire d'avoir un autre marqueur de sélection. Le marqueur de sélection utilisé pour la sélection des transformants, est la blasticidine. Le début de cette construction a été réalisé mais elle n'a pu être complètement caractérisée (fonctionnellement) faute de temps.

2. Réseau non-régulé

Aucun des tests effectués n'apporte à lui seul l'évidence définitive quant à l'action inhibitrice d'E1AFL (une fois activé par le promoteur *Cup1*) sur l'activité de DB-RB,. Cependant, on observe une même "tendance" récurrente dans les 3 tests de transactivation. Tous trois indiquent que :

- E1AFL affecte le phénotype du contrôle 2 en le "révertant" (partiellement à tout le moins),
- E1ACR2 n'affecte pas le phénotype du contrôle 2.

C'est parce que trois tests différents pointent dans la même direction que nous pensons que l'activité de DB-RB est effectivement inhibée par E1AFL, mais à un niveau faible.

Il semble que dans la cinétique également, un effet du cuivre sur la transcription d'E1A (via le promoteur *Cup1*) soit mis en évidence. Cependant, la concentration en cuivre dans le milieu ne semble pas optimale pour observer une diminution drastique voire totale de la coloration. De plus, les tests \square -Gal sont d'habitude réalisés après avoir laissé les levures 24 heures sur un milieu riche (YEPD). Or, le test a été effectué (en partie) à des temps beaucoup plus courts (0, 2, et 5 heures); il est normal, dans ces conditions, de ne pas observer le signal attendu.

3. Réseau oscillant

La construction du réseau oscillant n'a été qu'abordée. Grâce au réseau non-régulé, nous savons qu'en présence de cuivre, l'activité de DB-RB est inhibée par E1AFL. Cette information est importante pour pouvoir, dans le réseau oscillant, "synchroniser" les levures dans un seul état fixe du réseau.

3.1. Reconstruction du réseau

Pour reconstruire ce réseau, il a fallu transformer des levures par un vecteur d'expression qui porte *Gall-E1A*. La stratégie, pour observer les oscillations, est de mettre en plus dans cette souche le gène E1A sous le contrôle du promoteur *Cup1*. De cette manière, sur un milieu supplémenté en cuivre, il est possible de synchroniser les cellules dans un état précis de la boucle de rétroaction, c'est-à-dire de faire en sorte que E1A inhibe de manière permanente l'activité de DB-RB. Ensuite, si le cuivre est retiré du milieu, on s'attend à ce que les cellules, de l'entiereté de la population, activent simultanément et de manière semblable la boucle de rétrocontrôle. On s'attend donc à observer une oscillation synchrone au niveau de toute la population.

Faute de temps, cette opération de synchronisation n'a pas été faite. Les levures n'ont été transformées que par *Gall-E1AFL* ou par *Gall-E1ACR2*. et les tests (de croissance et β -Gal) ont été tentés.

3.2. Caractérisation phénotypique : résultats des tests de croissance et des tests β -Gal

Des conditions identiques aux tests du réseau non-régulé ont été utilisées ; à savoir, sur différents milieux déplétés en certains acides aminés, en présence de différentes concentration de cuivre et en présence ou absence de zéocine.

	Leu- Trp-	YEPD	Leu- Trp- + 5-FOA	Leu- Trp- His- + 3AT
Zéocine (150 μg/ml)	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre
Sans Zéocine	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre	0.0 mM Cuivre 0.0625 mM Cuivre 0.125 mM Cuivre 0.25 mM Cuivre 0.75 mM Cuivre

Aucun des tests effectués (ni les tests de croissance, ni les tests β -Gal) ne produisent de résultats différents des contrôles. Il est probable que la cause de l'absence d'effet soit simplement l'échelle à laquelle l'observation est réalisée. En effet, si l'ensemble des cellules n'est pas dans le même état du module, il est raisonnable de penser que ce qui est observé n'est qu'une "moyenne" des cellules individuelles ce qui estompe les oscillations éventuelles. De plus, les tests de croissance (pousse ou ne pousse pas) ne sont sans doute pas les plus adéquats

pour suivre la dynamique du module. Il serait plus intéressant de suivre en direct et en continu de l'expression individuelle de rapporteur dans une cellule à la fois.

Quelques réflexions rétro- et prospectives à propos de ce travail

1. Expériences à finaliser à propos des deux réseaux testés

Nous avons tenté de reconstruire, *in vivo*, deux réseaux de régulation répondant à des interactions protéines-protéines faisant intervenir RB, E2F et E1A. Nous avons d'abord construit un outil, pZACR42, qui nous permet, grâce à la technologie du MultiSite Gateway™, de combiner à volonté un promoteur à une ORF d'intérêt.

Le premier réseau (dit non-régulé) voit, l'activité de DB-RB complètement éteinte par E1A (sous le contrôle du promoteur *Cup1*) activé en permanence. Le second réseau (dit oscillant) doit montrer une activité de DB-RB inhibée de manière périodique par E1A (sous le contrôle du promoteur *Gall*) qui est lui-même placé sous le contrôle de DB-RB et donc présent un comportement dit oscillant.

Nous avons mis en évidence, dans le cadre du module non-régulé, que l'activité de DB-RB est, au moins en partie, inhibée par E1AFL.

Cependant, afin de mener le projet à terme, il est nécessaire et possible de réaliser quelques expériences complémentaires avant de continuer l'étude de ce module de régulation avec pour but d'éprouver l'approche adoptée c'est-à-dire d'écarter autant que possible les hypothèses.

1.1. Le module-pilote non-régulé

Accroissement du signal

La faible diminution de la croissance (sur un milieu sans histidine et supplémenté en 3AT) et de la décoloration (en test β -Gal) des levures transformées par le vecteur d'expression qui porte *Cup1*-E1AFL peut être due à une faible activation de la transcription de E1A. En d'autres mots, il est possible que la concentration en nitrate de cuivre utilisée ne soit pas optimale. Nous avons testé la gamme de concentration allant de 0.0 mM à 0.75 mM. Il serait intéressant d'élargir la gamme testée vers les plus hautes concentrations.

Contrôles

Lorsque les tests de transactivation sont effectués sur un milieu supplémenté en zéocine, nous n'avons aucun contrôle. Les souches utilisées comme contrôle phénotypique dans le double-hybride ne possèdent pas la résistance à la zéocine. Il sera donc nécessaire de transformer ces cinq souches, mais principalement les contrôles 1 et 2, par le vecteur de destination pZACR42 pour leur conférer cette résistance. De la sorte, il sera possible de comparer strictement la croissance et la couleur des levures testées (en tests de croissance et test β -Gal) par rapport aux contrôle 1 (contrôle positif) et contrôle 2 (contrôle négatif).

Quantification de la diminution de signal observée en test β -Gal

Les tests β -Gal que nous avons réalisés sont des tests qualitatifs puisqu'on juge visuellement la coloration ou qu'on utilise une méthode de densitométrie d'images, elle-même non quantitative. Pour un dosage quantitatif, il est nécessaire de réaliser un test β -Gal en culture liquide avec comme substrat chromogène le CPRG ou l'ONPG (plutôt que le X-Gal). Nous avons effectué une tentative qui a échoué mais faute de temps, elle n'a pu être renouvelée.

Observation d'une cellule individuelle

Le comportement du réseau au sein d'une cellule peut être masqué par un "nivellement" au niveau de la population entière. Le "vidéotracking" (suivi par vidéo d'une cellule via un logiciel informatique) d'une cellule individuelle doit pour cette raison être réalisé.

Choix du readout

Lorsque le réseau de type oscillant sera construit et son comportement éprouvé, il sera intéressant de suivre en "live" la variation continue d'expression du "readout". En effet, on s'attend à ce que les oscillations du réseau soient relativement rapides (de l'ordre de l'heure). Il est donc plus aisé et plus sensible de suivre l'évolution du "readout" (donc du module) que de devoir effectuer des mesures discontinues. C'est pour cette raison que nous comptons intégrer au génome de la levure un gène de GFP modifié [41] (fluorescence accrue, biais des codons adapté et court temps de demi-vie). L'intégration du gène de la GFP se fera au site du gène *URA3*. Les intégrants seront sélectionnés sur un milieu contenant du 5-FOA : s'il y a eu intégration, les cellules pousseront sur un tel milieu (sélection sur la perte du marqueur d'auxotrophie *URA3*).

1.2. Le réseau oscillant

Une fois la reconstruction du réseau oscillant réalisée, des tests seront entrepris. Il est possible qu'aucun signal ne soit observé (c'est le cas d'un certain nombre d'études publiées sur le sujet). Plusieurs explications et plusieurs solutions sont à évoquer.

Les oscillations sont des changements d'un état à un autre. L'état alternativement ON puis OFF du "readout" révèle celui du module. Pour expliquer l'absence de signal variable, on peut évoquer par exemple que les variations d'un état à un autre soient tellement rapides qu'elles en deviennent non-détectables via les changements d'état du "readout". C'est-à-dire que le cycle synthèse-dégradation de la GFP sera si court, que sa concentration restera pratiquement constante. Pour cette raison, il peut être intéressant d'accentuer l'amplitude des oscillations par ajout d'une drogue qui ralentit la traduction (cycloheximide).

Egalement important à considérer sont la vitesse de dégradation et la vitesse de synthèse des différents éléments du module. Celles-ci devront sans doute être adaptées si par exemple un des éléments se trouve dans un état stable (par exemple actif en permanence).

De plus, il sera nécessaire d' "équilibrer les deux jambes" de la boucle de rétrocontrôle [24, 25] [26] . Si une des deux relations entre éléments (soit l'inhibition de DB-RB par E1A, soit l'activation de E1A par DB-RB/AD-E2F) est trop forte, aucun signal "readout" ne sera observé. En perspectives moins immédiate, on peut envisager de remplacer E1A-FL par les véritables acteurs moléculaires de la régulation du cycle cellulaire (CDK, Cyclines, ...) et observer si effectivement le comportement du module est oscillatoire.

2. Perspectives de reconstruction à plus grande échelle

A long terme, on peut aussi s'adresser à d'autres modules. Ainsi, à partir de certaines données d'interactomique obtenue par double-hybride (par exemple carte de l'immunité de *C. elegans*, non-publiée) [42] [43] et en connaissant le sens des interactions des éléments les uns sur les autres, on peut envisager d'explorer ces cartes en suivant la méthode adoptée c'est-à-dire en reconnaissant, puis en reconstruisant certains modules ou sous-réseaux fonctionnels du réseau global. On conçoit que ces modules "simples", bien que complexes à réaliser et à

étudier, ne sont qu'une étape de l'approche synthétique de la biologie des systèmes. Il conviendra de greffer des modules les uns sur les autres et d'en observer le comportement. Techniquement, pour cela, il faudra probablement intégrer toutes les constructions dans le génome de la levure. En effet, le problème des marqueurs de sélection, en nombre limité, est critique. On peut aussi envisager d'avoir recours à des YACs (Yeast Artificial Chromosome). Le test ultérieur de ce genre de reconstruction sera de "complémenter" un mutant pour une machinerie de régulation spécifique avec son correspondant c'est-à-dire le réseau artificiel reconstruit par ingénierie génétique.

3. La question de la pertinence de l'approche que nous avons suivie

La "biologie des systèmes" emprunte ses concepts et bagages théoriques aux travaux d'autres disciplines (mathématiques, informatique, physique,...) et prétend aborder de manière plus globale l'interprétation des données de la biologie moléculaire et de la physiogénomique. Si sa nécessité pour appréhender la complexité des signalisations protéiques et génétiques, dont dépendent des phénomènes tels que la différenciation, l'adaptation, le développement (mais également la régulation dans les écosystèmes) ne fait depuis longtemps plus de doute, les outils pour l'aborder ont fait défaut jusqu'il y a peu.

A côté des approches analytiques (p. ex le décryptage de cartes d'interactions par l'étude de la "disruption" de certains nœuds ou lien, soit *in silico*, soit *in vivo*), on peut en effet envisager aussi des approches synthétiques dès lors que les outils de clonage, tels le MultiSite Gateway™, le permettent. C'est, à l'heure actuelle, dans cette perspective analytique et synthétique que se situe la recherche d'un petit nombre d'équipes de biologistes-physiciens.

Les travaux de Davidson (sur l'oursin) [44], de Barabaci-Albert et de Albert [45] (sur la drosophile) et ceux de Alon (sur *E. coli*) [46] [47, 48] [49] visent tous à la description des logiques de circuits et motifs à valeur prédictive. Ainsi, la logique booléenne "appliquée" sur les données biologiques relatives au développement primordial de la drosophile a permis à Réka Albert de proposer un modèle logique (Albert *et al.*, en cours de publication) qui simule p.ex l'apparition des métamères de l'embryon. Le "bombing" de certains éléments du circuit reproduit en simulation les modifications du pattern de développement de mutants correspondant *in vivo*. Le formalisme logique est évidemment indispensable pour expérimenter, à bon escient, dans de tels réseaux complexes qui se soustraient complètement à la compréhension intuitive.

Pour ce qui est de l'approche synthétique, on l'a vu, les expériences sont très peu nombreuses. Une seule a, à notre connaissance, interrogé un circuit de contrôle basé sur des interactions protéine-protéine [50, 51]. Les résultats sont extrêmement encourageants dans la mesure où le comportement du circuit fait découvrir des propriétés émergentes qui n'étaient pas soupçonnées et que les données offrent une valeur heuristique certaine pour l'expérimentateur. L'importance d'une interrogation plus systématique des réseaux de régulation protéines-protéines ne peut être sous-estimée. Considérant les cartes interactomiques déjà en notre possession, par exemple pour le modèle *C. elegans* (Vidal), on réalise combien c'est probablement à ce niveau que doit être recherchée l'explication du paradoxe du nombre de gènes, considéré comme faible, par rapport aux phénotypes si variés de l'organisme.

1. Techniques de base de Biologie moléculaire

1.1. Utilisation des enzymes de restriction

- 5 µl NEBuffer (10X) en fonction de l'enzyme (2, 3 ou 4),
- ER (Enzyme de Restriction) 2 µl + BSA 5µl (10X) si nécessaire,
- 5 µl de plasmide (de Miniprep, 2.5-5 µg/µl),
- Porter à volume final (50 µl) avec de l'eau distillée.

Incuber une heure dans un bain-marie à 37°C (seule exception : ApaI, 25°C – sur la paillasse)

Toutes les enzymes de restrictions viennent de *New England Biolab*.

Nom	No du tampon	BSA	Nbre U/ µl
<i>ApaI</i>	4	Oui	10
<i>BamHI</i>	2	Oui	20
<i>BspEI</i>	3	Non	10
<i>HindIII</i>	2	Non	20
<i>KpnI</i>	2	Oui	15
<i>NotI</i>	3	Oui	10
<i>SacII</i>	4	Non	10
<i>SnaBI</i>	4	Oui	5
<i>XhoI</i>	2	Oui	20

1.2. Fill in par le fragment Klenow de la DNA polymérase I des brins d'ADN protrudants

- 45 µl de plasmide restreint,
- 10 µl NEBuffer 4 (10X),
- dNTP's (100µM), chacun 2µl (total = 8µl),
- 2 µl Klenow (10 U) (*NEB*),
- Porter à volume final (100µl) avec de l'eau distillée(35 µl).

La réaction se déroule à température ambiante, sur la paillasse pendant 15 minutes.

1.3. Déphosphorylation des extrémités 5'-phosphate de brins d'ADN

- 45 µl de plasmide,
- 10 µl de NEBuffer 3 (10X),
- 0.25 µl de CIP Alkaline Phosphatase (2.5 U) (*NEB*),
- Porter à volume final (100 µl) avec de l'eau.

La réaction se déroule pendant deux heures dans un bain-marie à 37°C.

1.4. Ligation d'un vecteur et d'un insert

- Vecteur et Insert dans un rapport 1 : 10
- 2 µl de Tampon (10X) T4 DNA Ligase
- 1 µl de T4 DNA Ligase (400 U) (*NEB*)
- Porter à volume final (20 µl) avec de l'eau.

Réaction réalisée sur la nuit à température ambiante (sur la paillasse), volume final jamais supérieur à 20 µl.

1.5. Réaction en Chaîne de la Polymérase (PCR)

1.5.1. Principe

Cette technique est utilisée pour amplifier rapidement un fragment d'ADN d'intérêt. Les séquences bordant le fragment doivent être connues pour pouvoir designer des amorces oligonucléotidiques (encore appelées primers) qui leur sont complémentaires. La matrice d'ADN est **dénaturée** par chauffage en présence d'un excès d'oligonucléotides et de désoxyribonucléotides. Le mélange est ensuite soumis à une diminution de la température (en fonction des amorces), qui permet **l'hybridation spécifique** des amorces. La séquence comprise entre les deux amorces est **amplifiée** par une ADN Polymérase thermorésistante (Platinum Taq DNA Polymérase HiFi - *Invitrogen*) à une température optimale pour le fonctionnement de la polymérase. Le cycle de dénaturation/hybridation/élongation est répété 28 fois.

Le produit PCR (amplicon) peut être utilisé pour le séquençage, pour insertion dans vecteur ou encore pour l'analyse diagnostic de clones.

1.5.2. Protocole de PCR (pour un volume final de 50 µl)

Réactifs	Volume pour une réaction PCR
Tampon 10X de la Taq DNA Polymérase	5 µl
dNTP's – 100 µM (pour chaque)	0,1 µl (pour chaque dNTP)
MgSO ₄ (50 mM)	2 µl
Primer 1 (50µM)	1 µl
Primer 2 (50 µM)	1µl
Platinum Taq DNA Polymérase HiFi	0,2µl
Eau distillée et autoclavée	Porter à 48 µl
Template	2 µl

1.5.3. Programme PCR

Température	Durée
1- 94°C	2'
2- 94°C	45 secondes
3- 56°C	1'
4- 68°C	2' (1' par kb)
5- Retour à l'étape 2	28 fois
6- 68°C	5'
7- 10°C	Indéfiniment

1.6. Electrophorèse en gel d'agarose

1.6.1. Principe

Il s'agit d'une méthode de séparation de fragments d'ADN, dans un champ électrique, en fonction de leur taille. Il est possible d'estimer la taille des fragments grâce à un marqueur de poids moléculaire qui migre simultanément et dont les tailles des différents fragments sont connues. La révélation de la présence d'ADN est réalisée grâce au bromure d'éthidium qui rend l'ADN visible aux UVs en s'intercalant. La vitesse de migration dépend à la fois de la taille du fragment (inversement proportionnelle à la taille), du pourcentage d'agarose du gel (au plus le gel est concentré – fort pourcentage, au mieux les grands fragments sont séparés), de la conformation de l'ADN (l'ADN super enroulé migre plus rapidement que l'ADN linéaire) et de l'intensité du champ électrique.

1.6.2. Réalisation d'un gel d'agarose 1%

- Dissoudre par chauffage 1 g d'agarose dans 100 ml de tampon TBE 1X,
- Ajouter 5 µl de bromure d'éthidium (10 mg/ml) pour 250 ml de gel,
- Une fois que la température du flacon permet de le tenir en main, couler dans un support adéquat muni d'un peigne,
- Charger les échantillons auxquels a été ajouté au moins un tiers de tampon de chargement,
- Marqueur de poids moléculaire : 1kb Plus DNA Ladder (*Invitrogen*) 10 µl dilué (32 µl de MW – 1 µg/µl – 150 µl de tampon orange, 150 µl d'eau distillée).

Tampon de chargement (pour 100 ml)

- Glycérol 60%
- EDTA 25 mM
- 100 mM Tris-HCl, pH 7.5
- Pigment orange

Tampon TBE 10X (Tris-Borate-EDTA)

- Tris-Base 890 mM
- Acide borique 890 mM
- EDTA 19 mM
- Ajustez le pH à 8.3

1.7. Purification sur gel : QIAquick Gel Extraction Kit Protocol (avec une microcentrifugeuse)

Ce protocole est prévu pour extraire et purifier de l'ADN de 70 bp à 10 kb de gels d'agarose standards. Jusqu'à 400 mg d'agarose peuvent être purifiés par colonne.

Remarques :

- La couleur du tampon QG indique que le pH est ≤ 7.5 .
 - Ajouter de l'éthanol (96-100 %) au tampon PE avant l'utilisation (voir l'étiquette du flacon pour le volume).
 - De l'isopropanol (100%) ainsi qu'un bloc chauffant ou un bain-marie à 50°C sont nécessaires.
 - Toutes les étapes de centrifugation sont réalisées à environ 13000 rpm dans une microcentrifugeuse.
 - De l'acétate de sodium 3 M, pH 5.0, peut être nécessaire.
1. **Exciser les fragments d'ADN d'intérêt du gel d'agarose au moyen d'un scalpel propre.**
Minimiser la taille de la tranche de gel en enlevant l'agarose en trop.
 2. **Peser le morceau de gel. Ajouter 3 volumes de tampon QG à un volume de gel (100 mg ~ 100 µl).**
Par exemple, ajouter 300 µl de tampon QG pour chaque 100 mg de gel. La quantité maximale de gel par colonne est de 400 mg ; pour les morceaux >400 mg, utiliser plus d'une colonne.
 3. **Incuber à 50°C pendant 10 minutes (ou jusqu'à ce que le morceau de gel soit complètement dissous). Pour faciliter la dissolution de l'agarose, vortexer le tube toutes les 2-3 minutes pendant l'incubation.**
IMPORTANT : solubiliser complètement l'agarose.
 4. **Après dissolution complète, vérifier que la couleur de la solution est jaune (comme celle du tampon QG sans agarose dissous).**
Si la couleur du mélange est orange ou violette, ajouter 10 µl d'acétate de sodium 3 M, pH 5.0 et mélanger. La couleur du mix deviendra jaune (indicateur de pH, càd pH adéquat)
 5. **Ajouter un volume d'isopropanol à l'échantillon et mélanger.**
Par exemple, si le morceau de gel pèse 100 mg, ajouter 100 µl d'isopropanol. Ne pas centrifuger l'échantillon à ce stade.
 6. **Placer une colonne QIAquick spin dans un tube de 2 ml fourni.**
Cette colonne se compose d'une galette d'hydroxyapatite de calcium qui lie l'ADN.
 7. **Pour lier l'ADN, charger une colonne QIAquick avec l'échantillon, et centrifuger pendant une minute.**
 8. **Décanner l'éluat et replacer la colonne QIAquick dans le même tube.**

9. **Pour laver, ajouter 0.75 ml de tampon PE sur la colonne et centrifuger pendant une minute.**
Remarque : si l'ADN doit être utilisé pour des applications sensibles aux sels, telles que ligation à bouts-francs ou encore séquençage direct, laisser sur la paillasse 2-5 minutes après addition du tampon PE avant de centrifuger.
10. **Décanner l'éluat et centrifuger la colonne QIAquick pendant une minute supplémentaire à vitesse maximum.**
IMPORTANT : L'éthanol résiduel du tampon PE ne sera complètement enlevé que si l'éluat est décanter avant cette étape supplémentaire de centrifugation.
11. **Placer la colonne QIAquick dans un eppendorf propre de 1.5 ml.**
12. **Pour éluer l'ADN, ajouter 50 µl d'eau distillée au centre de la membrane de la colonne QIAquick et centrifuger pendant une minute à vitesse maximum. La diminution de la concentration en sels permet le détachement de l'ADN de son support.**
IMPORTANT : S'assurer que l'eau est ajoutée directement sur la membrane de la colonne QIAquick pour une élution complète de l'ADN. Le volume élué, en moyenne, est de 48 µl à partir d'un volume de 50 µl.

1.8. Purification sur colonne : QIAquick PCR Purification Kit Protocol (avec une microcentrifugeuse)

Ce protocole est prévu pour purifier des fragments d'ADN simple ou double brins de réactions PCR ou d'autres réactions enzymatiques.

Remarques :

- Ajouter de l'éthanol (96-100 %) au tampon PE avant l'utilisation (voir l'étiquette du flacon pour le volume).
 - Toutes les étapes de centrifugation sont réalisées à environ 13000 rpm dans une microcentrifugeuse.
1. **Ajouter 5 volumes de tampon PB à un volume d'échantillon PCR et mélanger.**
 Par exemple, ajouter 500 µl de tampon PB à 100 µl d'échantillon PCR.
 2. **Placer une colonne QIAquick spin dans un tube collecteur de 2ml fourni.**
 Cette colonne se compose d'une galette d'hydroxyapatite de calcium qui lie l'ADN.
 3. **Pour lier l'ADN, charger l'échantillon sur la colonne et centrifuger pendant 30-60 secondes.**
 4. **Décanner l'éluat. Replacer la colonne QIAquick dans le même tube.**
 5. **Pour laver, ajouter 0.75 ml de tampon PE à la colonne QIAquick et centrifuger 30-60 secondes.**
 6. **Décanner l'éluat et replacer la colonne QIAquick dans le même tube. Centrifuger pendant une minute supplémentaire à vitesse maximum.**
IMPORTANT : L'éthanol résiduel du tampon PE ne sera complètement enlevé que si l'éluat est décanter avant cette étape supplémentaire de centrifugation.
 7. **Placer la colonne QIAquick dans un tube eppendorf de 1.5 ml propre.**
 8. **Pour éluer l'ADN, ajouter 50 µl d'eau distillée au centre de la membrane de la colonne QIAquick et centrifuger pendant une minute. La diminution de la concentration en sels permet le détachement de l'ADN de son support.**
IMPORTANT : S'assurer que l'eau est ajoutée directement sur la membrane de la colonne QIAquick pour une élution complète de l'ADN. Le volume élué, en moyenne, est de 48 µl à partir d'un volume de 50 µl.

1.9. Minipréparation (Miniprep) : QIAprep Spin Miniprep Kit Protocol (avec utilisation d'une microcentrifugeuse)

Ce protocole est prévu pour purifier jusqu'à 20 µg d'ADN plasmidique à partir d'une culture liquide (sur la nuit) de 1-5 ml de *E. coli* dans du milieu LB (Luria-Bertani). Prendre 1.5 ml de culture et transférer dans un eppendorf propre. Centrifuger dans une microcentrifugeuse à vitesse maximum pendant un minute. Décanner le surnageant, puis commencer le protocole.

1. **Suspendre le culot de bactéries dans 250 µl de tampon P1 et transférer dans un tube pour microcentrifugeuse.**
S'assurer que la RNase A a été ajoutée au tampon P1. Aucun agrégat de cellules ne devrait être visible.
2. **Ajouter 250 µl de tampon P2 et agiter doucement le tube 4-6 fois pour mélanger.**
Mélanger doucement par inversion. Ne pas vortexer, ceci provoque un cisaillement de l'ADN génomique. Si nécessaire, continuer à agiter jusqu'à ce que la solution devienne visqueuse et légèrement claire. Ne pas laisser la réaction de lyse se dérouler plus de 5 minutes.
3. **Ajouter 350 µl de tampon N3 et agiter immédiatement doucement le tube 4-6 fois.**
Pour éviter une précipitation locale, mélanger la solution doucement mais complètement, immédiatement après addition du tampon N3. La solution devrait devenir trouble.
4. **Centrifuger pendant 10 minutes à vitesse maximum dans une microcentrifugeuse.**
Un culot blanc se formera.
5. **Transférer le surnageant de l'étape 4 à une colonne QIAprep par décantation ou pipetage.**
6. **Centrifuger pendant 30-60 secondes. Décanter l'éluat.**
7. **Laver la colonne QIAprep en ajoutant 0.75 ml de tampon PE et centrifuger pendant 30-60 secondes.**
8. **Décanter l'éluat, et centrifuger pendant une minute supplémentaire pour enlever les résidus de tampon de lavage.**
IMPORTANT : Les restes de tampon de lavage ne seront complètement enlever que si l'éluat est décanter avant cette centrifugation supplémentaire. L'éthanol restant du tampon PE peut inhiber les réactions enzymatiques ultérieures.
9. **Placer la colonne QIAprep dans un eppendorf de 1.5 ml propre. Pour éluer l'ADN, ajouter 50 µl d'eau distillée au centre de la colonne QIAprep, laisser sur la paillasse une minute, et centrifuger pendant une minute.**

2. TOPO Clonage (*Invitrogen*)

2.1. Principe

Le système de clonage Blunt®TOPO®PCR fournit une stratégie de clonage hautement efficace (qui prend cinq minutes) pour l'insertion directe de produits PCR à bouts francs générés par une ADN Polymérase proofreading dans un vecteur.

Le vecteur (pCR®-Blunt II-TOPO®) est fourni linéarisé. La DNA topoisomérase I du virus de la vaccine est liée de manière covalente aux extrémités 3' de chaque brins (qui est appelé vecteur "TOPO-activé").

La topoisomérase I du virus de la vaccine lie de l'ADN double brin à des sites spécifiques et le clive après la séquence 5'-CCCTT (coupure monobrin). L'énergie du lien phosphodiester clivé est conservée par la formation d'un lien covalent entre le phosphate 3' du brin clivé et un résidu tyrosyl (Tyr-274) de la topoisomérase I. Le lien phospho-tyrosyl entre l'ADN et l'enzyme peut être attaqué par la suite par l'extrémité 5' hydroxyl du brin original clivé, réaction inverse qui relâche la topoisomérase I. Le Clonage TOPO® exploite cette réaction pour cloner de manière efficace des produits PCR.

De plus, le vecteur pCR®-Blunt II-TOPO® permet une sélection directe des recombinants via la disruption du gène toxique de *E. coli*, *ccdB*. Le vecteur contient le gène *ccdB* fusionné à l'extrémité C-terminale du fragment LacZα. La ligation d'un produit PCR à bouts francs disrupte l'expression du gène de fusion *lacZ*α-*ccdB* ce qui permet la croissance uniquement des recombinants positifs après transformation. Les cellules qui contiennent un vecteur non-recombinant sont tuées une fois étalées.

2.2. Réalisation pratique

Le tableau ci-dessous décrit comment mettre en place une réaction de TOPO Clonage pour une transformation dans des cellules chimocompétentes ou électrocompétentes.

Remarque : La couleur bleue de la solution du vecteur pCR®-Blunt II-TOPO® est normale et est utilisée pour visualiser la solution.

Réactifs	<i>E. coli</i> chémocompétentes
Produit PCR	0.5 à 4 µl
Solution saline	1 µl
Solution saline diluée	--
Eau stérile	Ajouter pour porter à un volume final de 5 µl
Vecteur TOPO®	1 µl
Volume final	6 µl

- Mélanger prudemment et incubé 5 minutes à température ambiante (22-23°C).
- Pour des fragments PCR de grandes tailles (> 1kb) une incubation de plus de 5' peut donner plus de colonies après transformation.
- Mettre 2 µl du mélange de ligation-TOPO dans 100 µl de bactéries TOPO thermocompétentes. Ne pas mélanger par pipetage.
- Incuber sur glace pendant 5-30'.
- Choquer les bactéries pendant 30 secondes dans un bain-marie à 42°C.
- Transférer immédiatement sur glace, ajouter 250 µl de milieu SOC à température ambiante.
- Bien fermer les tubes et agiter horizontalement (200 rpm) à 37°C pendant 60'.
- Etaler sur milieu sélectif (LB + kanamycine 50 µg/ml), laisser pousser sur la nuit et lire le résultat le lendemain.

3. Réaction de MultiSite Gateway™

3.1. Réaction de BP Clonage

Réactif	Volume
Tampon BP 5X	2 µl
pDONR 201 (100 ng/□µl)	1 µl
Produit PCR	□□µl
TE 1X 10 mM, pH 8.0	3 µl
BP Clonase™	2 µl

Sortir la BP Clonase™ au dernier moment. Laisser décongeler sur glace (ou sur carboglace), prélever le volume nécessaire et remettre au congélateur (-80°C) immédiatement après utilisation. Pipeter pour mélanger. Veiller néanmoins à éviter la formation de bulles (tampon et solution de BP Clonase™ très visqueux).

Laisser la réaction se réaliser sur la nuit.

Prendre 2 µl du mélange de BP Clonage pour transformer 20 µl de bactéries (DH5α).

3.2. Réaction de LR Clonage

Réactif	Volume
Tampon LR 5X	2 µl
Vecteurs d'entrée (Promoterome et Orféome)	3 µl
Vecteur de destination restreint (pZACR42) – 100 ng	□□µl
LR Clonase™	2 µl

Sortir la LR Clonase™ au dernier moment. Laisser décongeler sur glace (ou sur carboglace), prélever le volume nécessaire et remettre au congélateur (-80°C) immédiatement après utilisation. Pipeter pour mélanger. Veiller néanmoins à éviter la formation de bulles (tampon et solution de LR Clonase™ très visqueux).

Laisser la réaction se réaliser sur la nuit.

Pour arrêter la réaction de clonage, ajouter 1 µl de protéinase K (2 mg/ml dans de l'eau), incubé 10' dans un bain-marie à 37°C.

Prendre 2 µl du mélange de LR Clonage pour transformer 20 µl de bactéries (DH5α).

4. Transformations

La transformation est une technique qui permet de faire rentrer de l'ADN exogène (plasmide) dans une cellule, bactérie ou levure.

4.1. Transformations bactériennes

- Placer 100 µl de bactéries thermocompétentes dans un eppendorf sur glace
- Une fois les bactéries décongelées, ajouter 5 µl du mélange de ligation
- Laisser sur glace pendant 30 minutes
- Choc thermique pendant 45 secondes dans un bain-marie à 42°C
- Immédiatement après le choc, remettre sur glace pendant 5 minutes
- Ajouter 2 ml de milieu SOC liquide
- Incuber une heure dans une pièce à 37°C en agitant (200 rpm)
- Centrifuger à 25°C, 2000 rpm pendant 5 minutes
- Décantier 1.8 ml de surnageant et resuspendre le culot de bactéries dans les 200 µl restant.
- Étaler les bactéries sur milieu sélectif approprié et laisser pousser sur la nuit à 37°C

4.2. Transformation de levures

1. Faire croître les levures sur la nuit à température ambiante, sous agitation (250 rpm) dans du milieu YEPD liquide. Utiliser 0.5 ml de culture par transformation.
2. Centrifuger 5' à 2000 rpm et laver avec du LiAc/TE (100 µl par transformation).
3. Centrifuger 5' à 2000 rpm et resuspendre dans du LiAc/TE (20 µl de par transformation).
4. Répartir les levures dans une plaque 96 puits, 20 µl par puits.
5. Ajouter 5 µl de vecteur d'expression (d'une Miniprep) à chaque puits
6. Faire un tube de PEG qui contient par réaction :
 - a. 100 µl de PEG/TE/LiAc
 - b. 2 µl d'ADN de sperme de saumon
7. Distribuer 100 µl de cette solution dans chaque puits. Mélanger par pipetage.
8. Incuber 30' à 30°C, puis 15' dans un bain-marie à 42°C.
9. Centrifuger 5' à 2000 rpm puis ôter le PEG.

Facultatif

10. Ajouter 100 µl de milieu YEPD liquide.
11. Incuber 60' à 30°C.
12. Centrifuger 5' à 2000 rpm puis ôter le YEPD.
13. Ajouter 100 µl d'eau distillée et autoclavée à chaque puits, ensuite retirer 90 µl et resuspendre dans les 10 µl restant (laisser moins de 10 µl dans le puits rend la resuspension difficile).
14. Déposé 4-6 µl sur un milieu approprié laisser pousser à 30°C.

TE/LiAc

- 10 mM Tris-HCl (pH 8.0),
- 1 mM EDTA,
- 0.1 M Acétate de Lithium (LiAc)

TE/LiAc/PEG

- TE/LiAc
- 44% polyéthylèneglycol 3350 (PEG, ultrapure, Sigma Catalog No. P-4338)

Salmon Sperm DNA (Sigma Catalog No. D-9156)

- Décongeler le tube d'ADN de sperme de saumon et porter un bain-marie à 100°C (eau bouillante),
- Mettre le tube d'ADN de sperme de saumon au bain-marie pendant 10',
- Immédiatement après, le mettre sur glace.

5. Souches et plasmides

5.1. Souches

5.1.1. Souches bactériennes (*E. coli*)

Souche	Génotype
DH5 alpha	F-, phi80dlacZDelta M15, Delta(lacZYA-argF), U169, deoR, recA1, hadR17(rk-,mk+),gal-, phoA, supE44, Lambda-, thi-1, gyrA96m relA1
DB3.1	F- gyrA462 endA- delta(sr1-recA) mcrB mrr hsdS20(rB-, mB-) supE44 ara14 galK2 lacY1 proA2 rpsL20(Smr) xyl5 lambda- leu mtl1
TOP10™	F- mcrA D(mrr-hsdRMS-mcrBC) f80lacZDM15 DlacX74 deoR recA1 araD139 D(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG

5.1.2. Souche de levure

Souche	Génotype
MaV203	<i>Mata leu2-3, 112 trp1-901 his3 Δ200 ade2-101 gal4 Δ gal80 Δ SPAL10::URA3 GAL1::lacZ GAL1::HIS3@LYS2 can1^R cyh2^R</i>

5.2. Plasmides (*Invitrogen*)

5.2.1. pPC86

Il s'agit du plasmide à partir duquel les constructions moléculaires ont été réalisées. Il porte un gène de résistance à l'ampicilline qui permet la sélection en bactéries.

5.2.2. pPICZalphaA

Plasmide à partir duquel la cassette zéocine [pTEF1-PEM7-*Sh Ble-CYC/ITT*] a été amplifiée.

5.2.3. pCR®-BluntII-TOPO®

Plasmide qui a permis le TOPO-Clonage de la cassette zéocine et duquel provient l'insert zéocine (excisé par restriction), cloné dans pACR42.

5.2.4. pDNR 201

Plasmide donneur de l'ORFéome qui porte un gène de résistance à la kanamycine qui permet la sélection des transformants. Après réaction de BP Clonage, on parle de vecteur d'entrée-ORFéome.

5.2.5. pDNR P4-P1R

Plasmide donneur du Promoterome qui porte un gène de résistance à la kanamycine qui permet la sélection des transformants. Après réaction de LR Clonage, on parle de vecteur d'entrée-promoterome.

6. Milieux

6.1. Milieux bactériens

6.1.1. Milieu SOB (pour 1 litre)

Réactifs	Quantité
2 % (w/v) Bactotryptone	20 g
0.5 % (w/v) Extrait de levure	5 g
10 mM NaCl	0.5844 g
2.5 mM KCl	0.1864 g
10 mM MgCl ₂	2.033 g
10 mM MgSO ₄	2.466 g

Ajouter un litre d'eau distillée.

Ajuster le pH entre 6.7 et 7.0 avec du KOH 5N.

Le SOC est un milieu SOB additionné d'une solution de glucose 20 mM (à partir d'un stock 40%, càd 2.22 M).

Glucose 40% (w/v)

Dissoudre 400 g de glucose dans un volume final de 1 litre.

Autoclaver la solution 45 minutes.

6.1.2. Milieu LB solide (Luria-Bertani)

- 10 g de bactotryptone
- 5 g d'extrait de levure
- 10 g de NaCl
- 1 litre d'eau
- Ajouter 15 g d'agar avant d'autoclaver. Autoclaver 45°C.
- Laisser refroidir dans un bain-marie à 55°C jusqu'à emploi
- Ajouter antibiotique si nécessaire puis couler dans des boîtes de Pétri stériles de 10 cm.

6.1.3. Milieu LB liquide (Luria-Bertani)

- 10 g de bactotryptone
- 5 g d'extrait de levure
- 10 g de NaCl
- 1 litre d'eau
- Autoclaver 45°C

Conserver à température ambiante.

Aux milieux LB solides comme liquides, il est possible d'ajouter des antibiotiques qui permettent une sélection de bactéries transformantes.

- Ampicilline 50 µg/ml
- Chloramphénicol 50 µg/ml
- Kanamycine 50 µg/ml

6.1.3. Milieu LB low-salt solide pour zéocine

- 10 g de bactotryptone
- 5 g d'extrait de levure
- 5 g de NaCl
- Ajouter 950 ml d'eau puis ajuster le pH à 7.5 avec du NaOH 1N puis porter le volume à un litre
- Autoclaver en cycle liquide pendant 45 minutes
- Laisser le milieu refroidir jusqu'à 55°C au moins puis ajouter la zéocine à la concentration voulue
- Couler le milieu dans des boîtes de pétri stérile de 10 cm.

6.2. Milieux levures

6.2.1. Milieu non-sélectif

Préparation de 2 litres de milieu solide

1. Mélanger 20 g d'extrait de levure (YE) et 40 g de bacto-peptone à 950 ml d'eau dans une flasque de deux litres et ajouter un barreau magnétique.
2. Ajouter 40 g d'agar à 950 ml d'eau dans une autre flasque de 2 litres.
3. Autoclaver pendant 45 minutes.
4. Mélanger l'agar à la «flasque-milieu».
5. Garder dans un bain-marie à 55°C jusqu'à utilisation.
6. Avant de couler le milieu dans les boîtes, ajouter 100 ml de glucose 40%.
7. Couler le milieu dans des boîtes de Pétri stériles de 15 cm.

Préparation de 2 litres de milieu liquide

1. Mélanger 20 g d'extrait de levure, 40 g de bacto-peptone à 1900 ml d'eau.
2. Autoclaver pendant 45 minutes.
3. Conserver à température ambiante.
4. Ajouter du glucose avant utilisation (50 ml de glucose 40% par litre).

6.2.2. Milieu sélectif

Préparation de 4 litres de milieu solide contenant un milieu synthétique complet (Sc) auquel il manque certains acides aminés ou nucléotides

1. Placer un barreau magnétique dans deux flasques appelées «flasques-milieu»
2. Mélanger 5.2 g de poudre d'acides aminés (-Trp, - His, - Leu, - Ura) (voir ci-dessous), 6.8 g de YNB (Yeast Nitrogen Base) et 20 g de sulfate d'ammonium.
3. Dissoudre dans 1900 ml d'eau et ajouter un barreau magnétique.
4. Ajuster le pH à 5.9 par addition de quelques gouttes de NaOH 10 M.
5. Ajouter 40 g d'agar à 900 ml d'eau dans deux autres erlenmeyer appelés « flasque-agar ».
6. Autoclaver les quatre flasques pendant 45 minutes.

Transférer le contenu de chaque flasque-agar dans une flasque-milieu.

Acides aminés

Pour préparer la poudre d'acides aminés (voir ci-dessus)

1. Mélanger 6 g de chacun des acides aminés suivant : alanine, arginine, acide aspartique, cystéine, acide glutamique, glutamine, glycine, isoleucine, lysine, méthionine, phénylalanine, proline, sérine, thréonine, tyrosine et valine.
2. Ajouter 6 g de sulfate d'adénine.

Remarque : le tryptophane, l'histidine, la leucine et l'uracile doivent être omis.

Les solutions concentrées (voir ci-dessus) sont préparées comme suit : 100 mM Histidine-HCl, 20 mM Uracile, 100 mM Leucine, et 40 mM Tryptophane. Ces solutions stocks concentrées peuvent être conservées à température ambiante, exception faite du tryptophane, qui doit être conservé dans le noir et à 4°C.

Milieu sélectif avec zéocine

Le pH doit être ajuster à 6.5 au lieu de 5.9. Autrement la zéocine est inactive.

7. Test β -Gal

1. Une solution stock de tampon Z est préparée en mélangeant:
 - 16.1 g $\text{Na}_2\text{HPO}_4 \cdot 7 \text{ H}_2\text{O}$ (ou 8.52 g anhydreux),
 - 5.5 g $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (ou 4.8 g anhydreux),
 - 0.75 g KCl,
 - 0.246 g $\text{MgSO}_4 \cdot 7 \text{ H}_2\text{O}$ (ou 0.12 g anhydreux) ,
 - dans un litre d'eau stérile et ajuster le pH à 7.0.
2. Répliquer les boîtes de levures sur des boîtes YEPD sur lesquelles se trouve un filtre de nitrocellulose (Osmonics, Catalog No. WP4HY13750).
3. Incuber sur la nuit à 30°C.
4. Pour chaque boîte, mettre deux filtres de papier Whatman (Whatman, Catalog No. 1454125) dans une boîte de Pétri de 15-cm vide.
5. Pour chaque boîte, utiliser 6 ml de tampon Z, auquel ont été ajoutés 11% μl 2-mercaptoéthanol et 100 μl de X-Gal 4% (5-bromo-4-chloro-3-indoyl- β -D-galactopyranoside) (dissous dans du diméthylformamide) (étape réalisée sous hotte)
6. Verser environ 200 ml d'azote liquide dans un bac à glace, mettre sous hotte.
7. Transférez le tampon Z dans les boîtes de Pétri contenant les papiers Whatman, s'assurez que l'entièreté du papier est mouillée et veillez à éliminer toutes les bulles d'air avec une pince à épiler.
8. Avec la pince à épiler, saisir le filtre de nitrocellulose qui porte les patches de levures et le plonger pendant 10 secondes dans l'azote liquide.
9. Mettre prudemment le filtre sur les papiers Whatman et retirer toutes les bulles d'air.
10. Incuber à 37°C , vérifier régulièrement et si nécessaire prendre des photos.
11. Incuber à 37°C sur la nuit puis lire le résultat.

BIBLIOGRAPHIE

1. Tegner, J., et al., *Reverse engineering gene networks: Integrating genetic perturbations with dynamical modeling*. Proc Natl Acad Sci U S A, 2003. **100**(10): p. 5944-9.
2. McAdams, H.H. and A. Arkin, *Towards a circuit engineering discipline*. Curr Biol, 2000. **10**(8): p. R318-20.
3. Jacob, F., Monod, J., *On the regulation of Gene Activity*. Cold Spring Harb Symp Quant Biol, 1961. **26**: p. 193-211.
4. Jacob, F., Monod, J., *Genetic regulatory mechanisms in the synthesis of proteins*. J. Mol. Biol., 1961. **3**: p. 318-356.
5. Monod, J., Jacob, F., *General Conclusions: Teleonomic Mechanisms in Cellular Metabolism, Growth and Differentiation*. Cold Spring Harb Symp Quant Biol, 1961. **26**: p. 389-401.
6. Shmulevich, I., et al., *Probabilistic Boolean Networks: a rule-based uncertainty model for gene regulatory networks*. Bioinformatics, 2002. **18**(2): p. 261-74.
7. Smolen, P., D.A. Baxter, and J.H. Byrne, *Mathematical modeling of gene networks*. Neuron, 2000. **26**(3): p. 567-80.
8. Shmulevich, I. and W. Zhang, *Binary analysis and optimization-based normalization of gene expression data*. Bioinformatics, 2002. **18**(4): p. 555-65.
9. Palsson, B., *The challenges of in silico biology*. Nat Biotechnol, 2000. **18**(11): p. 1147-50.
10. Ideker, T., T. Galitski, and L. Hood, *A new approach to decoding life: systems biology*. Annu Rev Genomics Hum Genet, 2001. **2**: p. 343-72.
11. Ideker, T., et al., *Integrated genomic and proteomic analyses of a systematically perturbed metabolic network*. Science, 2001. **292**(5518): p. 929-34.
12. Hartwell, L.H., et al., *From molecular to modular cell biology*. Nature, 1999. **402**(6761 Suppl): p. C47-52.
13. Bhalla, U.S. and R. Iyengar, *Functional modules in biological signalling networks*. Novartis Found Symp, 2001. **239**: p. 4-13; discussion 13-5, 45-51.
14. Neves, S.R. and R. Iyengar, *Modeling of signaling networks*. Bioessays, 2002. **24**(12): p. 1110-7.
15. Reboul, J., et al., *C. elegans ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression*. Nat Genet, 2003. **34**(1): p. 35-41.
16. Greenspan, R.J., *The flexible genome*. Nat Rev Genet, 2001. **2**(5): p. 383-7.
17. Rives, A.W. and T. Galitski, *Modular organization of cellular networks*. Proc Natl Acad Sci U S A, 2003. **100**(3): p. 1128-33.
18. Thomas, R., *De la molécule au mammifère*. 1983: p. 353-380.
19. Thomas, R., *Logical analysis of systems comprising feedback loops*. J Theor Biol, 1978. **73**(4): p. 631-56.
20. Thomas, R., *Boolean formalization of genetic control circuits*. J Theor Biol, 1973. **42**(3): p. 563-85.
21. Thieffry, D. and R. Thomas, *Dynamical behaviour of biological regulatory networks--II. Immunity control in bacteriophage lambda*. Bull Math Biol, 1995. **57**(2): p. 277-97.
22. Thomas, R., D. Thieffry, and M. Kaufman, *Dynamical behaviour of biological regulatory networks--I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state*. Bull Math Biol, 1995. **57**(2): p. 247-76.

23. Thomas, R. and M. Kaufman, *Multistationarity, the basis of cell differentiation and memory. II. Logical analysis of regulatory networks in terms of feedback circuits*. Chaos, 2001. **11**(1): p. 180-195.
24. Ferrell, J.E. and W. Xiong, *Bistability in cell signaling: How to make continuous processes discontinuous, and reversible processes irreversible*. Chaos, 2001. **11**(1): p. 227-236.
25. Ferrell, J.E., Jr., *Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability*. Curr Opin Cell Biol, 2002. **14**(2): p. 140-8.
26. Pomerening, J.R., E.D. Sontag, and J.E. Ferrell, Jr., *Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2*. Nat Cell Biol, 2003. **5**(4): p. 346-51.
27. Hasty, J., D. McMillen, and J.J. Collins, *Engineered gene circuits*. Nature, 2002. **420**(6912): p. 224-30.
28. Hasty, J., et al., *Synthetic gene network for entraining and amplifying cellular oscillations*. Phys Rev Lett, 2002. **88**(14): p. 148101.
29. Pardee, A., Jacob, F., Monod, J., *The Genetic Control and Cytoplasmic Expression of "Inducibility" in the Synthesis of b-galactosidase*. J. Mol. Biol., 1959. **1**: p. 165-178.
30. Friedman, D.I. and D.L. Court, *Bacteriophage lambda: alive and well and still doing its thing*. Curr Opin Microbiol, 2001. **4**(2): p. 201-7.
31. Kitano, H., *Computational systems biology*. Nature, 2002. **420**(6912): p. 206-10.
32. Kitano, H., *[Introductions to systems biology]*. Tanpakushitsu Kakusan Koso, 2003. **48**(7): p. 789-93.
33. Walhout, A.J., et al., *GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes*. Methods Enzymol, 2000. **328**: p. 575-92.
34. Vidal, M., et al., *Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system*. Proc Natl Acad Sci U S A, 1996. **93**(19): p. 10321-6.
35. Vidal, M., et al., *Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions*. Proc Natl Acad Sci U S A, 1996. **93**(19): p. 10315-20.
36. Helt, A.M. and D.A. Galloway, *Mechanisms by which DNA tumor virus oncoproteins target the Rb family of pocket proteins*. Carcinogenesis, 2003. **24**(2): p. 159-69.
37. Harlow, E., et al., *Association of adenovirus early-region 1A proteins with cellular polypeptides*. Mol Cell Biol, 1986. **6**(5): p. 1579-89.
38. Whyte, P., N.M. Williamson, and E. Harlow, *Cellular targets for transformation by the adenovirus E1A proteins*. Cell, 1989. **56**(1): p. 67-75.
39. Whyte, P., et al., *Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product*. Nature, 1988. **334**(6178): p. 124-9.
40. Giniger, E., S.M. Varnum, and M. Ptashne, *Specific DNA binding of GAL4, a positive regulatory protein of yeast*. Cell, 1985. **40**(4): p. 767-74.
41. Mateus, C. and S.V. Avery, *Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry*. Yeast, 2000. **16**(14): p. 1313-23.
42. Kaufman, M., J. Urbain, and R. Thomas, *Towards a logical analysis of the immune response*. J Theor Biol, 1985. **114**(4): p. 527-61.
43. Kaufman, M., F. Andris, and O. Leo, *A logical analysis of T cell activation and anergy*. Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3894-9.

44. Davidson, E.H., D.R. McClay, and L. Hood, *Regulatory gene networks and the properties of the developmental process*. Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1475-80.
45. Albert, R. and H.G. Othmer, *The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster*. J Theor Biol, 2003. **223**(1): p. 1-18.
46. Shen-Orr, S.S., et al., *Network motifs in the transcriptional regulation network of Escherichia coli*. Nat Genet, 2002. **31**(1): p. 64-8.
47. Alon, U., et al., *Response regulator output in bacterial chemotaxis*. Embo J, 1998. **17**(15): p. 4238-48.
48. Alon, U., et al., *Robustness in bacterial chemotaxis*. Nature, 1999. **397**(6715): p. 168-71.
49. Milo, R., et al., *Network motifs: simple building blocks of complex networks*. Science, 2002. **298**(5594): p. 824-7.
50. Becskei, A., B. Seraphin, and L. Serrano, *Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion*. Embo J, 2001. **20**(10): p. 2528-35.
51. Becskei, A. and L. Serrano, *Engineering stability in gene networks by autoregulation*. Nature, 2000. **405**(6786): p. 590-3.